



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2002, 20:55:58 ; Search time 1486.74 Seconds  
(without alignments)  
8651.612 Million cell updates/sec

Title: US-09-842-316-1

Perfect score: 1197

Sequence: 1 atggagtcgggctgctgcg.....cagaaccggctgcagactga 1197

Scoring table:

IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estfun:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estom:\*  
5: em\_estpl:\*  
6: em\_estba:\*  
7: em\_estov:\*  
8: em\_estov:\*  
9: em\_hic:\*  
10: gb\_est1:\*  
11: gb\_est2:\*  
12: gb\_hic:\*  
13: gb\_gss:\*  
14: em\_gss\_fun:\*  
15: em\_gss\_hum:\*  
16: em\_gss\_inv:\*  
17: em\_gss\_pin:\*  
18: em\_gss\_pro:\*  
19: em\_gss\_rod:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	568.8	47.5	757	11	BI181178
C 2	346.2	28.9	487	10	AW488801
C 3	326	27.2	725	11	BF383651
C 4	292.8	24.5	763	11	BI329992
C 5	291.8	24.4	295	11	BF353652
C 6	271.2	22.7	3015	12	AK004591
C 7	252	21.1	973	13	CNS02502
C 8	242.2	20.2	1126	11	BG682434
C 9	223.6	18.7	960	13	CNS01WAS
C 10	216.6	18.1	694	11	BI332488
C 11	216.6	18.1	793	11	BG744043
C 12	216.6	18.1	886	10	AL560216

13	215.2	18.0	907	10	AL560608
14	208	17.4	745	11	BF281109
15	197.2	16.5	695	11	BF281107
16	196.2	15.9	924	13	CNS02REF
17	175.4	15.0	347	10	AI317881
18	174.4	14.6	281	10	BB567592
19	174.4	14.6	460	10	BE655413
20	173.2	14.5	912	11	BG540337
21	168	14.0	535	11	BF188941
22	166.2	13.9	779	11	BG967782
23	155.6	13.3	707	10	AL543816
24	156.8	13.1	687	13	AZ733984
25	151.8	12.7	700	11	BG65008
26	148.2	12.4	673	11	BG921319
27	146.4	11.7	640	10	BE289740
28	135.8	11.3	938	10	AL552902
29	133.4	11.1	615	11	BF313287
30	133	11.1	616	11	BG611254
31	132.4	11.1	790	10	AU133824
32	131.8	11.0	680	10	AW077960
33	130.2	10.9	695	11	BE911051
34	129.4	10.8	879	10	AL553876
35	125.2	10.5	467	13	AQ388418
36	119.8	10.0	795	10	AL561703
37	118.8	9.9	467	10	AI218191
38	117	9.8	591	10	AW077347
39	116	9.7	790	11	BG685195
40	115.6	9.7	481	10	AI892564
41	114	9.5	715	10	AI588721
42	113.6	9.5	957	10	AL582021
43	110.8	9.3	532	10	AW182856
44	110	9.2	1068	13	CNS03YVB
45	109	9.1	810	10	AI958899

#### ALIGNMENTS

RESULT 1  
BI181178/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BI181178 757 bp mRNA EST 10-JUL-2001  
UNL-P-FN-ad-b-04-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone  
UNL-P-FN-ad-b-04-0-UNL 3', mRNA sequence.

BI181178

EST

GI:14655587

pig.

Sus scrofa

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

(bases 1 to 757)

Caetano, A.R., Johnson, R.K., and Pomp, D.

Generation and sequence characterization of a normalized cDNA

library from swine ovarian follicles

Unpublished (2001)

Contact: Pomp, D

Department of Animal Science

University of Nebraska, Lincoln

Lincoln, NE 68583-0908, USA

Tel: 402 472 6416

Fax: 402 472 6362

Email: dpomp@unl.edu

Oligo-dT track not found, Not 1 site shown in beginning of sequence

is likely internal to the message. The following repetitive

elements were found in this cDNA sequence: 495-552,

>GC-rich/low-complexity

Seq primer: M13 -29

POLYA-No.

Location/Qualifiers

1..757

/organism="Sus scrofa"

/strain="University of Nebraska, Lincoln Swine Selection

Lines"











200k21 of library G from Tetraodon nigroviridis, genomic survey sequence.	
ACCESSION	AL170173
VERSION	AL170173.1 GI:7808230
KEYWORDS	GSS; genome survey sequence.
SOURCE	Tetraodon nigroviridis.
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE	1 (bases 1 to 960)
AUTHORS	Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 960)
AUTHORS	Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 960)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <a href="http://www.genoscope.cns.fr/Tetraodon">http://www.genoscope.cns.fr/Tetraodon</a> .
FEATURES	Location/Qualifiers
source	1..960
	/organism="Tetraodon nigroviridis"
	/db_xref="taxon:99883"
	/clone="200K21"
	/clone_lib="G"
	/note="Genoscope sequence ID : C0AG200AF11LP1-end : 27"
BASE COUNT	195 a 268 c 233 g 261 t 3 others
ORIGIN	
Query Match	18.7%; Score 223.6; DB 13; Length 960;
Best Local Similarity	57.1%; Pred. No. 2e-30;
Matches 481; Conservative	3; Mismatches 347; Indels 11; Gaps
QY	40 gtcattgctctgcattacaactacacggcgaagctccgcggtgcgctaccagcgggt 99
Db	
130	gtccctagtgaaagcattacactacgacgtggcaagctacagaagacggctacc-----gt 183
QY	100 gcggcctgcgcgcgcgcgcgtggtgctgcggtgtgcccctcatctgctagaag 159
Db	
184	gaaggctcaaacggaggcattggccctcttgggtgctgcctgcctatgtgcttgag 243
QY	160 aatctagcgtgtgtgtgtgtgctgcgacgccacccgcttcacgctcccatgttctctg 219
Db	
244	aatgcgtgtggtgtctgtgagccatttgaggaaacaaagattccacattgcccattat 303
QY	220 ctctctggcagcctcagctgtcgattctgctgcgagcgccgcctacgcgcccaacatc 279
Db	
304	ctgttggaacaccttactttgtcgcgacctctgtagctgggattacattatgtgccaaatc 363
QY	280 ctactctggggccgctcaactgtgaactgtccccgcgctctgtgttcgacggaggga 339
Db	
364	atcatctgtcgacacctaactcgttcgaaactgcacacctctgattgtgttcacggaggga 423
QY	340 ggctcttcctggcactcaactggttcgctctgagctcctgcctcgcctcgctggagcgc 399
Db	
424	cgagcttcattacacttagccgcctctgttaattagctctgtgcccaattgcaattgaacga 483
QY	400 agcctcaccatggcgcagggggccgcgcgcctctccagtcgggggcgcgcctctgagcgc 459





```

Db 404 AGCTGCTCTTCACTGAGGGAGCGCTTTGCCACCATGGTGGGGCCGGTGGCCGAGAGC 463
QY 433 ---gtctccagtcggggcgacgctggtgagatggcagcggcgctggggcgctgctgctg 489
Db 464 GGGGCCACCAAGACAGCCGCGTCTACGGCTTCATGGGCTCTGCTGGCTGCTGCGCCGCG 523
QY 490 ctctcgggctccctcgccagcgctggctggaattgctgctggctggctggaactgtctcc 549
Db 524 CTGCTGGGATGCTGCTTTGCTGGGCTGGAACCTGCTGGCCCTTTGACCGCTGCTCC 583
QY 550 actgtcttgcgcgtctacgccaaggctacgtgctgtcttcttgcgtgctgcttctgctg 609
Db 584 AGCTTCTGCGCCCTCTACTCCAGCGCTACATCTCTTCTGCTGGTGATCTTCGCCGCG 643
QY 610 atctcgccactctgtgcactctacgctgagcgatctactgcaagtagcagcgaacg 669
Db 644 GTCTGGCCACCATATGGCGCTTATGGGCGCATCTTCGCGCTGGTGAGGCCAGCGGG 703
QY 670 c99 672
Db 704 CAG 706

RESULT 13
AL560608 907 bp mRNA EST 16-FEB-2001
LOCUS AL560608 LTI_NFL010_BC2 Homo sapiens cDNA clone CS0DL003YC21 5
DEFINITION prime, mRNA sequence.
ACCESSION AL560608
VERSION AL560608.1 GI:12907235
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 907)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. .907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DL003YC21"
/clone_lib="LTI_NFL010_BC2"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/vector="Vector: pCMVSPORT 6; Site: 1; NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 105 a 327 c 299 g 176 t
ORIGIN

Query Match 18.0%; Score 215.2; DB 10; Length 907;
Best Local Similarity 56.3%; Pred. No. 6.1e-29;
Matches 508; Conservative 0; Mismatches 353; Indels 41; Gaps 4;

QY 13 ctgctcgccggcgccggtgagcgagtgatctgcttctgattacactacacggcaag 72
Db 29 CAGCTGGCGCGCGGGGACAGCGCGCTCATTTGTTGCTACTACACCACTACCGGCGG 88

```

```

QY 73 ctccgcggtgctgctaccagccggtgctgctgctgctgctgctgctgctgctgctgctg 132
Db 89 CTGGCGGCGCGGGGGCC-GGAGGATGGCGGCTGGGGGCGCTTCGGGTG 147
QY 133 ggggtgagccttcagctgctagagaatctagcgtgtgtgtgtgtgtgtgtgtgtgtgtgt 192
Db 148 GCGCCAGCTGCTGGTGTGCTGGAGAACTTGTGCTGGCGGCGCATCACCAGCCAC 207
QY 193 cgcgcttcacgctcccatgttctgctgctgggagcctcactgctgctgctgctgctgctg 252
Db 208 ATCGGCTCGCGAGCGCTGGGTCTACTATTGCTTGGTGGTGAACATCAGCTGAGTGACTG 267
QY 253 gaaggcgccgctacgcccacacatctactgtctgctgctgctgctgctgctgctgctgct 312
Db 268 ACGGCGCGGCGCTTACTTGGCCAACTGCTGTGCTGGGGGCGCGCACCTTCGCTGTGGCG 327
QY 313 cgcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 372
Db 328 CCGCGCCAGTGGTCTCTAGGGGAGGCGCTTCTTACCGCGCGCTGGCGCGCTTCCACCTTC 387
QY 373 agcctcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 432
Db 388 AGCTGCTCTTCTACTGACAGGAGCGCTTTGCCACCATGGTGGCGCGCTGGCGGAGAGC 447
QY 433 ---gtctccagtcggggcgacgctgctgctgctgctgctgctgctgctgctgctgctg 489
Db 448 GGGGCGCCACCAAGACAGCGCGCTTACGGCTTCTTACGGCTTCTTCTGCTGCTGGCGCG 507
QY 490 ctctcgggctcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 549
Db 508 CTGCTGGGATGCTGCTTCTTGGGCTGGAACTGCTGCTGGCGCTTGGCGCTTGGACCGCTG 567
QY 550 actgtctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 609
Db 568 AGCTTCTGCGCCCTTACTTCCAAAGCGCTACATCTCTTCTGCTGGTGGTGTCTTCCCGCG 627
QY 610 atctcggcgctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 669
Db 628 GTCTGGCGCACCATCATGCTGGGCTCTTATGGGCGCATCTTCCGCGCTGGGCGCGCGGG 687
QY 670 cgcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 729
Db 688 CAGAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 726
QY 730 cgcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 789
Db 727 -----GCCTGCTGAAGACGCTGCTGATGATCTTCTGCTGGCGCTTCTTCTGCTGCTG 773
QY 790 tggggccccccttctctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 849
Db 774 TGGGCGCCACATCTTGGGCTGCTGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 833
QY 850 gtaacctctgagggc---gatacctctctgctgctgctgctgctgctgctgctgctgctgct 906
Db 834 GAGTACCTCGCGGCGCATGGACTGATGATCTGCTGGCGCTTCTTCTGCTGCTGCTGCTGCTG 893
QY 907 cc 908
Db 894 CC 895

RESULT 14
BF281109 745 bp mRNA EST 28-NOV-2000
LOCUS BF281109
DEFINITION EST445700 Rat Gene Index, normalized rat, Rattus norvegicus CDNA
ACCESSION Rattus norvegicus CDNA clone RG1AA02, mRNA sequence.
VERSION BF281109
KEYWORDS BF281109.1 GI:11212179
SOURCE EST.
ORGANISM Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

```







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2002, 20:58:03 ; Search time 1761.82 Seconds  
(without alignments)  
11208.366 Million cell updates/sec

Title: US-09-842-316-1  
Perfect score: 1197  
Sequence: 1 atggagtcgggctgctgctg.....cagaaccggctgcagactga 1197

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sy.\*
- 13: gb.un.\*
- 14: gb.vl.\*
- 15: em.ba.\*
- 16: em.fun.\*
- 17: em.hum.\*
- 18: em.in.\*
- 19: em.om.\*
- 20: em.or.\*
- 21: em.ov.\*
- 22: em.pat.\*
- 23: em.ph.\*
- 24: em.pl.\*
- 25: em.ro.\*
- 26: em.sts.\*
- 27: em.sy.\*
- 28: em.un.\*
- 29: em.vl.\*
- 30: em.htgo\_hum.\*
- 31: em.htgo\_inv.\*
- 32: em.htgo\_rod.\*
- 33: em.htg\_hum.\*
- 34: em.htg\_inv.\*
- 35: em.htg\_rod.\*
- 36: em.htg\_other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1197	100.0	1197	6	AX147832	AX147832 Sequence
2	1197	100.0	100680	9	AC011461	AC011461 Homo sapi
3	1195.4	99.9	1197	9	AF317676	AF317676 Homo sapi
4	1195	99.8	1198	6	AX138796	AX138796 Sequence
5	774	64.7	1203	10	AF233649	AF233649 Rattus no
6	768	64.2	206533	2	AC073749	AC073749 Mus muscu
7	766.4	64.0	2437	10	BC012232	BC012232 Mus muscu
8	764.4	63.9	2171	10	AF115249	AF115249 Rattus no
9	734.8	61.4	1165	10	AF327535	AF327535 Homo sapi
10	642.8	53.7	62147	2	AC026510	AC026510 Homo sapi
11	294.8	24.6	978	10	AY011730	AY011730 Lemur cat
12	289.8	24.2	978	10	AY011704	AY011704 Muscardin
13	277.2	23.2	977	4	AY011726	AY011726 Artibeus
14	275	23.0	978	4	AY011694	AY011694 Sorex ara
15	274.6	22.9	978	4	AY011717	AY011717 Ochotona
16	271.4	22.7	978	10	AY011710	AY011710 Dipodomys
17	271.2	22.7	1149	10	AF108019	AF108019 Mus muscu
18	269.6	22.5	1149	6	AX085541	AX085541 Sequence
19	269.6	22.5	7177	10	W6040811	U40811 Mus musculu
20	268.4	22.4	270	6	AX147764	AX147764 Sequence
21	268.2	22.4	978	10	AY011708	AY011708 Hystrix b
22	268	22.4	978	10	AY011714	AY011714 Myocastor
23	265.4	22.2	978	4	AY011693	AY011693 Condylura
24	261.6	21.9	2232	6	AR027719	U10303 Rattus norv
25	261.6	21.9	2232	6	I32245	AK027719 Sequence
26	261.2	21.8	978	10	AF011706	I32245 Sequence 3
27	261.2	21.8	978	10	AF164114	AF164114 Eugu rubr
28	259.2	21.7	1695	5	AF164114	AF164114 Eugu rubr
29	258.6	21.6	978	10	AY011703	AY011703 Castor ca
30	257.4	21.5	975	4	AY011739	AY011739 Leopardus
31	257.4	21.5	975	4	AY011740	AY011740 Panthera
32	257.4	21.5	978	4	AY011737	AY011737 Tapirus i
33	257.4	21.5	978	9	AY011718	AY011718 Cynoceph
34	256.8	21.5	978	4	AY011719	AY011719 Tupia mi
35	255.8	21.4	975	4	AY011738	AY011738 Felis cat
36	255.8	21.4	978	4	AY011736	AY011738 Caracotha
37	255.8	21.4	978	10	AY011715	AY011715 Agouti ta
38	254.8	21.3	978	4	AY011735	AY011735 Equus cab
39	254.2	21.2	978	4	AY011729	AY011729 Megaptera
40	253	21.1	978	10	AY011711	AY011711 Heterocep
41	252	21.1	1149	9	AF233365	AF233365 Homo sapi
42	252	21.1	77678	2	AL161741	AL161741 Homo sapi
43	252	21.1	132438	9	HS0575N6	AL109741 Human DNA
44	251.6	21.0	978	4	AY011742	AY011742 Manis pen
45	251	21.0	978	4	AY011732	AY011732 Trageleph

ALIGNMENTS

RESULT 1	AX147832	AX147832	1197 bp	DNA	PAT	08-JUN-2001
LOCUS	Sequence	77 from Patent	WO0136473.			
DEFINITION	AX147832					
ACCESSION	AX147832.1	GI:14346837				
VERSION						
KEYWORDS	human.					
SOURCE						
ORGANISM	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 1197)					
AUTHORS	Vogeli, G., Wood, L.S., Parodi, L.A., Hiebsch, R.R., Lind, P.,					
	Slignton, J., Schellin, K.A., Kaytes, P.S., Bannigan, C.M., Ruff, V.,					
	Sejltz, T. and Huff, R.M.					
TITLE	Novel g protein-coupled receptors					
JOURNAL	Patent: WO 0136473-A 77 25-MAY-2001;					
	PHARMACIA & UPJOHN COMPANY (US)					
FEATURES	Location/Qualifiers					
source	1. .1197					

/organism="Homo sapiens"

/db\_xref="taxon:9606"

BASE COUNT 139 a 449 c 389 g 220 t

ORIGIN

Query Match 100.0%; Score 1197; DB 6; Length 1197;  
Best Local Similarity 100.0%; Pred. No. 1.3e-167;  
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atcgagtcgggctgctgagcgccgctgagcgaggtcatctgctgattacaac 60  
Db 1 ATGGAGTCGGGGCTGCTGCGGCGCGCGCTGAGCGAGGTCTGCTGCTGATTACAA 60  
QY 61 tacacggcgaagctccgggtgagcgctaccagccgggtgagcggtgagcgccagcc 120  
Db 61 TACACGGCAAGCTCCGGGTGCGGTACACAGCGGTGCGGCTGCGCGCGACGCC 120  
QY 121 gtgggtgcttgggggtgctgagccttcatctgctgagaaatctagcgtgtgtggtg 180  
Db 121 GTGGGTGCTGCTGGGGGTGCGGCTTCTATGCTGTAGAGAACTAGCCGTGTGTGTG 180  
QY 181 ctggagacccacccgctcttcacgctcccatgttctctctgagcgagcctcagttg 240  
Db 181 CTCGAGCCCAACCGCGCTTCCAGCTCCCATGTTCTGCTCTGCGAGCCTCAGTTG 240  
QY 241 tgggtatgctggagcgccgctacgagccgcaacatctactgtcggggcgctcagc 300  
Db 241 TCGGATCTGCTGGAGCGCGCTAGCGCGCAACATCTACTGTGCGGGCGGCTCAGC 300  
QY 301 ctgaaactgtcccgccgctggttcgacggagggagggaggtctctgagcactact 360  
Db 301 CTGAAACTGTCCCGCGCTGCTGCTGCGAGGGAGGGGCTCTCTGCTGCACTCACT 360  
QY 361 gctgctgctgagcctcctggccatcgctgagcgagcagcctaccatggtcgagc 420  
Db 361 GCGTCCGTGCTGAGCTCTGGCCATCGCGCTGAGCGCAGCTCACCATGCGCGCAGG 420  
QY 421 gggccgcccgcctcagtcggggggcgacgctggcgatggcgagcgccgtggggc 480  
Db 421 GGGCCCGCGCGCTCTCAGTGGGGGGCGACGCTGGCGATGGCAGCGCGCGCTGGGG 480  
QY 481 gtgtgctgctcctgggctcctgacagcgctgggtggtgaaattgctgggtcgctgac 540  
Db 481 GTGTGCTGCTCCTCGGCTCTGCGAGCGCTGGGTGGAATGCTGGGTGCTGCTGAC 540  
QY 541 gcttgcctcactgtctgctgctgctgctgctgctgctgctgctgctgctgctg 600  
Db 541 GCTTGTCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600  
QY 601 ttctgagggcattcgtgctgctgctgctgctgctgctgctgctgctgctgctg 660  
Db 601 TTCTGAGGCATCTGCGCGCTATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
QY 661 gccacgctgggctgctgctgctgctgctgctgctgctgctgctgctgctgctg 720  
Db 661 GCCAACGCGCGCTGCTGCGCGCAGCGCGCGCTGCGCGCAGCTGCTGCGCGCGCG 720  
QY 721 cgtcgaagcgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 780  
Db 721 CGTGCAGCGCGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780  
QY 781 gtggcattgtggggcgccctcttctgctgctgctgctgctgctgctgctgctgctg 840  
Db 781 GTGGCATGTTGGGGCGCCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
QY 841 acctgtcttactctgagcgcgatcccttctgctgctgctgctgctgctgctgct 900  
Db 841 ACCTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
QY 901 ctgacaccccaatctacagcgtcacaacacccgagccttgcgcacgagcctgagc 960  
Db 901 CTGAAACCCATCATCTACAGCTTACCAACCGGAGCTTGGCGCCAGCGCTCTCTG 960

QY 961 gtctgtcggagcgcactcctcgcgcagagaccgagtggtgctccagcagtcggcgagc 1020  
Db 961 GTCTGTCTGGGAGCGCCACTCTCTGGGCGAGAGACCAGTGGCTCCAGCAGTCCGCGAGC 1020  
QY 1021 gcggtgaggtctccgggggctgctgctgctgctgctgctgctgctgctgctgctg 1080  
Db 1021 GCGGTGAGGCTTCGGGGGCGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080  
QY 1081 agcggtcggagcgtctatcgccccagcgagcggtgagcacacagcggtccacagc 1140  
Db 1081 AGCGGTCTGGAGCGCTCATCGCCCCAGCGCGAGCGGTGAGCACACAGCGGTCTCCACAGC 1140  
QY 1141 agccccggtgacccacagccgccccgagctgtggtatcagacacgctgacagctga 1197  
Db 1141 AGCCCCGCTGCACCCACAGCGCGCGGACTCTGTTATCAGAACCCGCTGCAGACTGA 1197

RESULT 2

AC011461 100680 bp DNA PRI 31-MAY-2001  
LOCUS Homo sapiens chromosome 19 clone CTC-429L19, complete sequence.  
DEFINITION AC011461  
ACCESSION AC011461  
VERSION AC011461.4 GI:14269680  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 100680)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 100680)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 100680)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (31-MAY-2001) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On May 31, 2001 this sequence version replaced gi:13752679.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 100% of Sequence;  
Estimated Total Number of Errors is 0.

STS Content:  
SHGC-53359 G36605  
SHGC-63493 G48302.

FEATURES

source

1. 100680  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTC-429L19"

BASE COUNT 24238 a 25942 c 26581 g 23919 t  
ORIGIN

Query Match 100.0%; Score 1197; DB 9; Length 100680;  
Best Local Similarity 100.0%; Pred. No. 1.8e-168;  
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 atcgagtcgggctgctgagcgccgctgagcgaggtcatctgctgattacaac 60  
Db 41760 ATGGAGTCGGGGCTGCTGCGGCGCGCGCTGAGCGAGGTCTGCTGCTGCTGCTGCTG 41819  
QY 61 tacacggcgaagctccggtgctgctgctgctgctgctgctgctgctgctgctgctg 120  
|||||

Db 41820 TACACGGCAAGCTCCGGGTGCGGTACACAGCGGGTGGCGGCTGCGCGCCAGCGCC 41879  
Qy 121 gtggtgctgctgagcggtgtgctgcttcacatgctgtagagaatcagcgtgtgtgtgtg 180  
Db 41880 GTGGTGTGCTTGGCGGTGCGGCTTCACTGCTGCTAGAGAACTACGCGGTGTGTGTGTG 41939  
Qy 181 ctggagcgcacccgcgctccacgctccacgctccacgctccacgctccacgctccacg 240  
Db 41940 CTGGAGCGCCACCGCGCTTCCAGCTCCCATGTTCTGCTCTCTGGCAGCCTCAGCTG 41999  
Qy 241 tcggatctgctgagcgccgctccacgctccacgctccacgctccacgctccacgctcc 300  
Db 42000 TCGGATCTGTGTCAGCGCGCGCTTCCAGCGCGCAACATCTACTGTCGGCGCGCTCAG 42059  
Qy 301 ctgaacgctgctcccgctgctgctgctgctgctgctgctgctgctgctgctgctgct 360  
Db 42060 CTGAACGCTGCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42119  
Qy 361 gctgctgctgagcctcctgctgctgctgctgctgctgctgctgctgctgctgctgct 420  
Db 42120 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42179  
Qy 421 gggcgcgccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 480  
Db 42180 GGGCGCGCGCGCTTCCAGTCCGGGGCGCACGTGGCGATGGCAGCGCGCGCTGGGG 42239  
Qy 481 gtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 540  
Db 42240 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 42299  
Qy 541 gctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 600  
Db 42300 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42359  
Qy 601 ttgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 660  
Db 42360 TTGCTGGGATCCTGGCGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42419  
Qy 661 gcaacgcgcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 720  
Db 42420 GCCAAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42479  
Qy 721 cgtcgcaagcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 780  
Db 42480 CGTGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42539  
Qy 781 gggatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 840  
Db 42540 GTGGATGTTGGGGCGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42599  
Qy 841 acctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 900  
Db 42600 ACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 42659  
Qy 901 ctgaaccccatcatatcacgctcaccaacgagcctgctgctgctgctgctgctgctgct 960  
Db 42660 CTGAACCCCATCATATCACGCTACCAACCGGACCTGGCCAGCGGCTCTGCGGCTG 42719  
Qy 961 gctgctgctgagcctcctgctgctgctgctgctgctgctgctgctgctgctgctgct 1020  
Db 42720 GTCTGCTGCGGACGCCATCTCTGCGGCGAGAGACCGAGTGGCTCCAGCAGTGGCG 42779  
Qy 1021 ggggctgagggcttcgggggctgctgctgctgctgctgctgctgctgctgctgctg 1080  
Db 42780 GCGGCTGAGGCTTCCGGGGGCTGCGGCGCTGCTGCTGCGGCGGCTTGTATGGAGCT 42839  
Qy 1081 agcgctgagcgctcctgctgctgctgctgctgctgctgctgctgctgctgctgctg 1140  
Db 42840 AGCGGCTGAGGCTCATCGCCAGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCG 42899  
Qy 1141 agccccgctgacccacagcgcccgagctctggtatcagaacccgctgagagctga 1197  
Db 42900 AGCCCCGTGACCCACAGCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42956

RESULT 3  
AF317676  
LOCUS  
DEFINITION  
Homo sapiens sphingosine 1-phosphate receptor Edg-8 gene, complete cds.  
AF317676  
AF317676.1 GI:11559845  
human.  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (31-OCT-2000) Pharmacology, University of Virginia School of Medicine, 1300 Jefferson Park Avenue, Charlottesville, VA 22908, USA  
FEATURES  
source  
1. 1197  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/note="similar to sequence presented in GenBank Accession Number AC011461"  
<1..>1197  
/product="sphingosine 1-phosphate receptor Edg-8"  
1. 1197  
/note="g protein-coupled receptor"  
/codon\_start=1  
/product="sphingosine 1-phosphate receptor Edg-8"  
/protein\_id="AAG38113.1"  
/db\_xref="GI:11559845"  
/translation="MESGLLRPAVSEVILVHNTGKLRGARYOPGAGLRADAVVCL AVCAFVLENLAVLLVGRPHRPFAPMFLLSLSDLLAGAAAYANILISGLPTLK LSPALFARGGVFTALTSVLSLAIALERSLTMARGPAPVSRGRTLAMAAWGV VLLGLLPALGNCLGRDASTVLPYAKAYLFCVLAFLAFVACWGPFLFLLLDVA VRANRLPARGTGTSTTRARKPRSLALLRSLVLAFLAFVACWGPFLFLLLDVA CPARTCPVLIQADPFLGLAMANSLLNPIYITLTNRDLRALRLVCCGRHSCGRDPSG SOOSASAAEASGGLRCLPPGLDGSFSGSERSPQRDGLDTSVSGTSGPAPTARTLV SEPAAD"  
BASE COUNT 139 a 449 c 390 g 219 t  
ORIGIN  
Query Match 99.9%; Score 1195.4; DB 9; Length 1197;  
Best Local Similarity 99.9%; Pred. No. 2.3e-167;  
Matches 1196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 atgagtgagggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 60  
Db 1 ATGAGTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60  
Qy 61 tacaccgcaagctcccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 120  
Db 61 TACACCGCAAGCTCCCGGTGCGGCTACCAAGCGGGTGGCGGCTGCGCGCCAGCGCC 120  
Qy 121 gtggtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 180  
Db 121 GTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
Qy 181 ctggagcgcacccgcgctccacgctccacgctccacgctccacgctccacgctccacgctg 240  
Db 181 CTCGGAGCGCCACCGCGCTTCCACGCTCCCATATGTTCTGCTGCTGCGGACGCTCAG 240  
Qy 241 tcggatctgctgagcgccgctcctcagcgcgcaacatcctactgctcggggcgctcagc 300















\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

```

1      655: contig of 655 bp in length
      666 765: gap of 100 bp
      766      1456: contig of 691 bp in length
      1457 1556: gap of 100 bp
      1557      2243: contig of 687 bp in length
      2244 2343: gap of 100 bp
      2344      3029: contig of 686 bp in length
      3030 3129: gap of 100 bp
      3130      3785: contig of 636 bp in length
      3786 3885: gap of 100 bp
      3886      4562: contig of 677 bp in length
      4563 4662: gap of 100 bp
      4663      5350: contig of 688 bp in length
      5351 5450: gap of 100 bp
      5451      6133: contig of 683 bp in length
      6134 6233: gap of 100 bp
      6234      6914: contig of 681 bp in length
      6915 7014: gap of 100 bp
      7015      7719: contig of 705 bp in length
      7720 7819: gap of 100 bp
      7820      8523: contig of 704 bp in length
      8524 8623: gap of 100 bp
      8624      9322: contig of 699 bp in length
      9323 9422: gap of 100 bp
      9423      10104: contig of 682 bp in length
      10105 10204: gap of 100 bp
      10205      10897: contig of 693 bp in length
      10898 10997: gap of 100 bp
      10998      11698: contig of 701 bp in length
      11699 11798: gap of 100 bp
      11799      12488: contig of 690 bp in length
      12489 12588: gap of 100 bp
      12589      13272: contig of 684 bp in length
      13273 13372: gap of 100 bp
      13374      14030: contig of 678 bp in length
      14031 14130: gap of 100 bp
      14131      14855: contig of 705 bp in length
      14856 14955: gap of 100 bp
      14956      15656: contig of 701 bp in length
      15657 15756: gap of 100 bp
      15757      16458: contig of 702 bp in length
      16459 16558: gap of 100 bp
      16559      17250: contig of 692 bp in length
      17251 17350: gap of 100 bp
      17351      18034: contig of 684 bp in length
      18035 18134: gap of 100 bp
      18135      18814: contig of 680 bp in length
      18815 18914: gap of 100 bp
      18915      19604: contig of 690 bp in length
      19605 19704: gap of 100 bp
      19705      20365: contig of 661 bp in length
      20366 20465: gap of 100 bp
      20466      21154: contig of 689 bp in length
      21155 21254: gap of 100 bp
      21255      21935: contig of 681 bp in length
      21936 22035: gap of 100 bp
      22036      22721: contig of 686 bp in length
      22722 22821: gap of 100 bp
      22822      23505: contig of 684 bp in length
      23506 23605: gap of 100 bp
      23606      24294: contig of 689 bp in length
      24295 24394: gap of 100 bp
      24395      25093: contig of 699 bp in length
      25094 25193: gap of 100 bp
      25194      25899: contig of 706 bp in length
```



[illegible]

[illegible]

RESULT	13
AY011726	
LOCUS	AY011726 977 bp DNA
DEFINITION	Artibeus jamaicensis EDGL (EDGL) gene, partial cds.
ACCESSION	AY011726
	MAM 14-FEB-2001

VERSION	AY011726.1	GI:12829917
KEYWORDS	Jamaican fruit-eating bat.	
SOURCE	Artibeus jamaicensis	
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Chiroptera; Microchiroptera; Phyllostomidae; Stenodermatinae; Artibeus.	
REFERENCE	1 (bases 1 to 977)	
AUTHORS	Murphy,W.J., Eisarik,E., Johnson,W.E., Zhang,Y.P., Ryder,O.A. and O'Brien,S.J.	
TITLE	Molecular phylogenetics and the origins of placental mammals	
JOURNAL	Nature 409 (6820), 614-618 (2001)	
MEDLINE	21082082	
REFERENCE	2 (bases 1 to 977)	
AUTHORS	Murphy,W.J., Eisarik,E., Johnson,W.E., Zhang,Y.P., Ryder,O.A. and O'Brien,S.J.	
TITLE	Direct Submission	
JOURNAL	Submitted (01-NOV-2000) Laboratory of Genomic Diversity, National Cancer Institute-FCRDC, P.O. Box B, Bldg. 560, Rm. 11-10,	
FEATURES	Frederick, MD 21702, USA	
Source	Location/Qualifiers	
mRNA	1..977	
gene	/organism="Artibeus jamaicensis"	
CDS	/db_xref="taxon:9417"	
	<1..>977	
	/gene="EDG1"	
	/product="EDG1"	
	<1..>977	
	/gene="EDG1"	
	<1..>977	
	/gene="EDG1"	
	/codon_start=3	
	/product="EDG1"	
	/protein_id="AAK01994.1"	
	/db_xref="GI:12829918"	
	/translation="ENGLTKTSVPFLICFSIIIVENIVLLTTTWTKKFKHRMYVFIC NLALSDLAGVAVINLPLSGANTVKLPQVHFLREGSMFVALSASVSFILLAIARIERY ITHWKLKNLSNCSRFELLISACWISLVGGPLMGNCVGLPSTVIPLPHKH YILFTCTAVVLALLSVIYICRIYSLVTRSRRLTFRNASKTSRSSEKSALLKTIV ILVSITFWBLLIDLDVGCKNTCSILFKAEYELVLAVLNSTNPITYLTATNKE MRRAFTIRILSCCKPRGEAGAKFRPIITAGVEFSRSKSNSSHPQKDGG"	
BASE COUNT	177 a 326 g 212 t	
ORIGIN		
Query Match	23.2%; Score 277.2; DB 4; Length 977;	
Best Local Similarity	56.9%; Pred. No. 5.5e-32;	
Matches 535; Conservative	0; Mismatches 393; Indels 12; Gaps 1;	
QY	102 cggcctgcgcgcgaacgccttggtgtgcctggcggtgtgcgccttcacgtcgtagagaa	161
Db	8 CGGCCTCAAGTGACCTCGGTGCTTCATTCTCATCTGCAGCTTCATCATCTCGAGAA	67
QY	162 ttatagccgttgttgggtgcgcgcgcgcacccgcgccttcacgcctcccattctctgct	221
Db	68 CATCTTTGCTGTGTGACCATTGGAAGAACCAAGAAGTTCACCGGCCCATGTACTATT	127
QY	222 cctgggcagcctcaactgttcgatctgtgcgcgcgcgcacctcacgcgcacaatcct	281
Db	128 CATCGCAACCTGCGCCCTCTCAGACCTGCTGCGGGGGTGGCCTACGTATCAACCTGCT	187
QY	282 actgtcggggccgcctcaacgtgaaactgtccccccgcgcctctggttcgcacgggagg	341
Db	188 TTCTCTCGGGGGCAACAACACCTACAAAGCTCACGCCCGCTGCAGTGGTTCCTCGGGGA	247
QY	342 cgctcttgtggcactcaactgccttcgtgtgcagtctctggccatcgctcgagcgacg	401
Db	248 CATGTTGTGTCCTGTGCGCGTCCGTGTTACGCTGCTGGCCATCGCCATCGAGCGCTA	307
QY	402 cctcaccatggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc	461
Db	308 CATCACCATGCTGAAGATGAAGCTGTACAAAGAACCAACAGCTGCCGCTCCTTCTGCT	367









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2002, 21:03:03 ; Search time 69.25 seconds  
(without alignments)  
3914.717 Million cell updates/sec

Title: US-09-842-316-1

Perfect score: 1197

Sequence: 1 atgagtcggggtgtgtcg.....cagaaccggtgcagactga 1197

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq.\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq.\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq.\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq.\*
- 5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq.\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	261.6	21.9	2232	1	US-08-196-989B-3
2	261.6	21.9	2232	2	US-08-760-936-3
3	248.4	20.8	1137	3	US-09-082-088-1
4	244	20.4	1637	3	US-08-852-824-3
5	227.4	19.0	1649	2	US-08-845-566-2
6	206.6	17.3	2754	1	US-08-196-989B-1
7	206.6	17.3	2754	2	US-08-760-936-1
8	148.6	12.4	1260	3	US-08-789-982-1
9	137.8	11.5	1761	3	US-08-861-747-1
10	137.8	11.5	1889	3	US-08-861-747-3
11	133.2	11.1	2250	3	US-08-763-938-1
12	98.8	8.3	1875	5	PCT-US93-10618-1
13	90	7.5	2185	2	US-08-467-948A-3
14	90	7.5	2185	3	US-08-467-947A-3
15	67.6	5.6	1417	4	US-09-199-737-3
16	67.2	5.6	1065	4	US-09-325-897-1
17	67	5.6	1356	3	US-08-997-803-13
18	66.4	5.5	1523	3	US-08-997-803-12
19	66.4	5.5	1621	1	US-08-722-001-13
20	66.4	5.5	2140	1	US-08-334-698-1
21	66.4	5.5	2140	1	US-08-228-932-1
22	66.4	5.5	2140	2	US-08-468-939-1
23	66.4	5.5	2140	2	US-08-406-855A-1
24	66.4	5.5	2140	2	US-08-722-190-1
25	66.4	5.5	2140	3	US-08-244-354-1
26	66.4	5.5	2140	3	US-09-206-899-1
27	66.4	5.5	2140	5	PCT-US95-04203-1

28	64.8	5.4	1776	1	US-08-722-001-29	Sequence 29, Appl
29	64	5.3	639	3	US-08-997-803-3	Sequence 3, Appl
30	63	5.3	2580	3	US-09-050-863-2	Sequence 2, Appl
31	63	5.3	5452	2	US-09-130-114-1	Sequence 1, Appl
32	63	5.3	9600	4	US-08-910-647-1	Sequence 1, Appl
33	63	5.3	10596	1	US-07-884-811-15	Sequence 15, Appl
34	63	5.3	10596	1	US-07-885-971-15	Sequence 15, Appl
35	63	5.3	10596	1	US-08-087-783A-15	Sequence 15, Appl
36	63	5.3	10596	1	US-08-194-088B-15	Sequence 15, Appl
37	63	5.3	10596	2	US-08-194-087-15	Sequence 15, Appl
38	63	5.3	10596	5	PCT-US93-04648-15	Sequence 15, Appl
39	62.2	5.2	1296	1	US-07-816-283-9	Sequence 9, Appl
40	62.2	5.2	1296	1	US-08-417-103-9	Sequence 9, Appl
41	59.4	5.0	2126	2	US-08-789-354-1	Sequence 1, Appl
42	59.4	5.0	2126	3	US-09-110-937-1	Sequence 1, Appl
43	59.4	5.0	2126	3	US-09-058-725B-1	Sequence 1, Appl
44	59.4	5.0	2126	3	US-09-232-857-1	Sequence 1, Appl
45	58.4	4.9	1394	2	US-08-068-729-3	Sequence 3, Appl

## ALIGNMENTS

RESULT 1  
US-08-196-989B-3  
; Sequence 3, Application US/08196989B  
; Patent No. 5585476  
; GENERAL INFORMATION:  
; APPLICANT: MacLennan, A. John  
; TITLE OF INVENTION: Molecular Cloning and Expression of  
; TITLE OF INVENTION: G-Protein Coupled Receptors  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/196,989B  
; FILING DATE: 15-FEB-1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lloyd, Jeff  
; REGISTRATION NUMBER: 35,589  
; REFERENCE/DOCKET NUMBER: MAC-100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2232 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 269..1420  
US-08-196-989B-3

Query Match 21.9%; Score 261.6; DB 1; Length 2232;  
Best Local Similarity 56.4%; Pred. NO. 1.7e-40;  
Matches 536; Conservative 0; Mismatches 399; Indels 15; Gaps 2;

OY 37 gaggtcatcgcttcgtacatacaccgcaagtcggtgcgtaccagccg 96







[illegible]

Qy	497	ggctctcgccagcgctgggctgaaattgctgctgggtcgctgagcattgtgtccaatgtct	556
Db	635	GTGGCTTGCCTATCTTGGCTGGAAATGTCTGGACATCTGGAGCTTGTCTCCACTGTGC	694
Qy	557	tgcgctctacgccaaagcctacgtgctcttctgcgtgctcgcttcgtaggcatactcgg	616
Db	695	TGCCCCCTATAGCTTAAGCACTATGTCTCTGCTGGTGCACCATCTTCTGTGTCATCTTAC	754
Qy	617	cgcgtatctgtgcactctacgcgcgcgcatctactgcaggtgtacgcccacgcgcgcgc	677
Db	755	TGSGCTATCGTGGCCCTTGTAGCTCCGAATCTACTTGTAGTCCGTCAAGCCATCGGGACG	814
Qy	677	tgcgggacgcggccgggagctgcggggaccacactgcagccgggcgctgcgaacgcgcgt	736
Db	815	TTGCTTGGT-----CTCTAGA	829
Qy	737	cgctggcctctgctgcgcagcctcagcgtggtgctctggtgctttgtggcatgttggggc	796
Db	830	CGCTGGCCCTGCTCAAGACAGATCACCATCGTACTGGGTGTGTTTTCATCATCTGTGTGCTGC	889
Qy	797	ccctctctctgctgctgttgcgcagctggcgtgccgggcgcacacttctctgtaactcc	856
Db	890	CGGCTTTTAGCATCCTTCTCTTAGACTACTACTGTCCGCTCGGGCTGTCTGTGCTTCT	949
Qy	857	tgcaggccgatacctctctggagctggccatgtggccaaactcactcttgaaccccaatcatct	916
Db	950	ACAAAGCCCATATTCTTTTGCCCTTGGCCACCCCTCAACTCTCTGCTCAACCCGTGCATCT	1009
Qy	917	acagcctcaccaaccgcagcactgcgccacgcgcctctgcgctggtctctgctgcggaaccc	976
Db	1010	ATACATGGCGTAGCCGGGACCTTCGGAGGGAGGTACTGAGGCCCTCTGCTGTCTGGCCGC	1069
Qy	977	a	977
Db	1070	A	1070

```

1  RESULT 7
2  US-08-760-936-1
3  ; Sequence 1, Application US/08760936
4  ; Patent No. 5856443
5  ; GENERAL INFORMATION:
6  ; APPLICANT: MacLennan, A. John
7  ; TITLE OF INVENTION: Molecular Cloning and Expression of
8  ; TITLE OF INVENTION: G-Protein Coupled Receptors
9  ; NUMBER OF SEQUENCES: 14
10 ; CORRESPONDENCE ADDRESS:
11 ; ADDRESSER: Saliwanchik, Lloyd & Saliwanchik
12 ; STREET: 2421 N.W. 41st Street, Suite A-1
13 ; CITY: Gainesville
14 ; STATE: FL
15 ; COUNTRY: US
16 ; ZIP: 32606
17 ; COMPUTER READABLE FORM:
18 ; MEDIUM TYPE: Floppy disk
19 ; COMPUTER: IBM PC compatible
20 ; OPERATING SYSTEM: PC-DOS/MS-DOS
21 ; SOFTWARE: Patentin Release #1.0, Version #1.25
22 ; CURRENT APPLICATION DATA:
23 ; APPLICATION NUMBER: US/08/760,936
24 ; FILING DATE: December 6, 1996
25 ; CLASSIFICATION: 536
26 ; ATTORNEY/AGENT INFORMATION:
27 ; NAME: Pace, Doran R.
28 ; REGISTRATION NUMBER: 38,261
29 ; REFERENCE/DOCKET NUMBER: MAC-100C1
30 ; TELECOMMUNICATION INFORMATION:
31 ; TELEPHONE: 352-375-8100
32 ; TELEFAX: 352-372-5800
33 ; INFORMATION FOR SEQ ID NO: 1:
34 ; SEQUENCE CHARACTERISTICS:
35 ; LENGTH: 2754 base pairs
36 ; TYPE: nucleic acid

```

RESULT 8  
US-08-789-982-1







Db 458 CCGCAGCCGCTTTACCTTGGGCTGTCTCCGCGAGGCTGTCTGGACACAAG 517  
QY 354 actactgctccgtgctgagctccctggccatcgctgagcagcctaccatggc 413  
Db 518 CTTACTCGCTGGTGGCCACACTGCTGGCCATCGCCGTGAACGACCGCAGTGTGAT 577  
QY 414 ggcagggggggccgcccgtctccagctgggggagcagcgtggcagcgccgccc 473  
Db 578 GCGCCGTACAGTTCACAGCCGCTGCCCCGCGGCTGCTATCTATTTGGCGCT 637  
QY 474 ctggggcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 533  
Db 638 GTGGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 697  
QY 534 cctgagcgttgcctccactgcttctgctgctgctgctgctgctgctgctgct 593  
Db 698 CTTGGACCGCTGCTCAGCGATGGACCCCTGCTCAGCGGCTCTATTTGGCGCT 757  
QY 594 gctcgcttgcgggctcctgctgctgctgctgctgctgctgctgctgctgct 653  
Db 758 TCTGTCGAGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 817  
QY 654 ggtacgcccgaacgcccgtgctgctgctgctgctgctgctgctgctgctgct 713  
Db 818 CTTGGCGGCGGAGTGCAGCGCATGGCAGAGCAT-----GTCAGCTGCCACCCCGCTAC 872  
QY 714 ccgggctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 773  
Db 873 CGAG-----AGACCACGCTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 919  
QY 774 ggccttggctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 833  
Db 920 GCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979  
QY 834 ggcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 893  
Db 980 TGAG---TCCTGCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1036  
QY 894 ctaacttgaacccatcatctacagctgctgctgctgctgctgctgctgctgct 953  
Db 1037 CTACTGGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096  
QY 954 ggcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 978  
Db 1097 CCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1121

## RESULT 11

US-08-763-938-1  
; Sequence 1, Application US/08763938  
; Patent No. 6140060  
; GENERAL INFORMATION:  
; APPLICANT: CHUN, Jerold J.M.  
; APPLICANT: HECHT, Jonathan H.  
; TITLE OF INVENTION: CLONED LYXOPHOSPHATIDIC ACID  
; TITLE OF INVENTION: RECEPTORS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikaido, Marmelstein, Murray and Oram LLP  
; STREET: 655 15th Street, N.W., Suite 330 - G St. Lobby  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/763,938  
; FILING DATE: 12-DEC-1996

CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: JAHNS, Kristina M.  
REGISTRATION NUMBER: 41,092  
REFERENCE/DOCKET NUMBER: P8074-6018  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638-5000  
TELEFAX: (202) 638-4810  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2250 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-763-938-1

Query Match 11.1%; Score 133.2; DB 3; Length 2250;  
Best Local Similarity 49.4%; Pred No. 7, 7e-17;  
Matches 422; Conservative 0; Mismatches 408; Indels 24; Gaps 2;

QY 125 tgtgctggcgtggcgtgctgctgctgctgctgctgctgctgctgctgctgctgct 184  
Db 391 TGGGCATCACTGTTGCTGTTTCATCATGTTGCCAATCTCTCTGCTCATGTGGCAATCT 450  
QY 185 gacgcccacgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 244  
Db 451 ACGTCAACCCCGCTTCCCAATTTCCCTTATTTACTTATGATGGCCAACTTGGCTGCTG 510  
QY 245 atctgctggcgtggcgtgctgctgctgctgctgctgctgctgctgctgctgctgct 304  
Db 511 ACTTCTCGCTGATGGCTACTTCTACTCATGTTCAATACAGGACCTTAATACCCGGA 570  
QY 305 aactgctccccgcgtgctgctgctgctgctgctgctgctgctgctgctgctgctgct 364  
Db 571 GACTGACTGTTAGCAGTGGCTTCCCTCCGCGAGGGCTCATTTGACACACGCTGACAGCTT 630  
QY 365 cgtgctgagcctcctggcctgctgctgctgctgctgctgctgctgctgctgctgctgct 424  
Db 631 CTGTGGCCAACTGCTGGCTATTGCTATCGAGAGGCACATCACGGTTTTCGCGATGACG 690  
QY 425 ccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 484  
Db 691 TCCATACAGCAATGAGCAACCGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 750  
QY 485 cgtgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 544  
Db 751 CCATTGTGATGGTGTATACCCAGTGTGGCTGGAACTGCTGTGATGATGATGATGAT 810  
QY 545 gctcactgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 604  
Db 811 GTTCAACATGGCACCCTCTACAGTGTACTTACTTACTTACTTACTTACTTACTTACTT 870  
QY 605 tgggcatcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 664  
Db 871 TGGTGAACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930  
QY 665 acgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 724  
Db 931 GGACTATG-----AGGATGCTCGGCATGTTCTGGACCCAGGAGGA 972  
QY 725 gcaagccgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 784  
Db 973 ATCGGACACCATGATGATGAGCCTTCTGAAGACTGTGGCTATGTTGGTGGCTTTATG 1032  
QY 785 catgttggggggccctcttctgctgctgctgctgctgctgctgctgctgctgctgct 844  
Db 1033 TCTGTGACTCCGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086  
QY 845 gctcgtactcctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 904  
Db 1087 GCGATCTCTGCTGCTATGAGAAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1146



STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,948A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04079  
FILING DATE: 30-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: STEFFE, ERIC K.  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2185 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 884...2062  
US-08-467-948A-3

Query Match 7.5%; Score 90; DB 2; Length 2185;  
Best Local Similarity 49.4%; Pred. No. 6.9e-09;  
Matches 333; Conservative 0; Mismatches 320; Indels 21; Gaps 3;  
QY 290 ggcgtcactgaaactgtcccccgcgtctgtgtgcagggagggagggcgtcttcg 349  
DB 1167 GTCCCGCACAGCCGACATTTACATTGAGGCTGTCTCTCGGCGAGGGCTGTGGACA 1226  
QY 350 tggcactcactgctgcgtgctgagcctcctgtgcccactgctgagcgcagcctcacc 409  
DB 1227 CAACCTCACTGCTCGGTGGCCACACTGCTGGCCATCGCCGTGGAGCGGACCGCAGTG 1286  
QY 410 tggcgcgcagggggggcgccgcgcgtctccagtcggggcgccagcgtgcgagcgccg 469  
DB 1287 TGATGCGGTGACAGTGCACAGCGCGCTGCCCTGGCGGCGGTGCTCATGTCTATTGGG 1346  
QY 470 cggcctggggcggtgctgctgctcctcctggtcctccagcgtggtggaattgcctgg 529  
DB 1347 GCGTGTGGTGGTGGCTTGGCCCTGGGCGCTGCTGCCCTGCCACATCCCTGGCAGCTCT 1406  
QY 530 gtgcctgagcgtgctgctcactgctgctgctgctgctgctgctgctgctgcttct 589  
DB 1407 GTCCCTGACCGCTCTCTCAGCAGTGGACCGCTGCTCAGCGGTCTCTATTGGCCGTCT 1466  
QY 590 ggtgctcgtccttcgtgggcatcctggcgctatctgtgcactctacgcgcgcactact 649  
DB 1467 GGGCTCTGTCGAGCTGCTGTGCTTCTTCCTGCTCATGTGTGTACACCGCATTTCT 1526  
QY 650 gccaggtagcgcacagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 709  
DB 1527 TCTACGTGCGCGGCGAGTGCAGCGCATG-----GCAGAGCATGTCAGTGCACCCCG 1581  
QY 710 cgcacccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 769  
DB 1582 CTACCGAGA-----GACCACGCTCAGCCTGGTCAAGACGTGTGTTCATCATCC 1628

QY 770 tctggcctttgtgcatgttggggccccccttcttctgctgctgtgtgtgctgacgtg 829  
DB 1629 TGGGGGCGTTGCTGCTGCTGACACAGCCAGGTGGTACTGCTCTGATGCTTTAG 1688  
QY 830 gccgcgcgcgcacactgtcctgtactcctgcagggcgatccctccttctggtgactg 889  
DB 1689 GCTGTGAG---TCTGCAATGCTTGGCGTTAGAAAAGTACTTCTTACTGTGGCCG 1745  
QY 890 ccaactcacttgaaccccatcatctacacgcctcaccacgcgcgcgcgcgcgcgc 949  
DB 1746 CAACCTCACTGGTCAATGCTGCTGTACTTCTGCGGAGATGCTGAGATGCGCGCACCT 1805  
QY 950 tctggcgcctggtc 963  
DB 1806 TCCGCGCCTCTC 1819  
RESULT 14  
US-08-467-947A-3  
Sequence 3, Application US/08467947A  
Patent No. 6030575  
GENERAL INFORMATION:  
APPLICANT: LI, YI  
APPLICANT: CAO, LIANG  
APPLICANT: NI, JIAN  
APPLICANT: GENTZ, REINER  
APPLICANT: BULT, CAROL J.  
APPLICANT: SUTTON III, GRANGER G.  
APPLICANT: ROSEN, CRAIG A.  
TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein  
TITLE OF INVENTION: Coupled Receptor GPR1  
NUMBER OF SEQUENCES: 30  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
STREET: 1100 NEW YORK AVE., NW, SUITE 600  
CITY: WASHINGTON  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,947A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04079  
FILING DATE: 30-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: STEFFE, ERIC K.  
REGISTRATION NUMBER: 36,688  
REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-371-2600  
TELEFAX: 202-371-2540  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2185 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 884...2062  
US-08-467-947A-3

Query Match 7.5%; Score 90; DB 3; Length 2185;  
Best Local Similarity 49.4%; Pred. No. 6.9e-09;

Matches 333; Conservative 0; Mismatches 320; Indels 21; Gaps 3;

QY 290 ggcgcctcagcgtgaactgtcccccgcgtctgttgcacgagggagggcgtcttcg 349  
 Db 1167 GTCCCGCACAGCCGACCTTACCTTGGGCTGGTCTTCTCGGCGAGGCTTGTGACA 1226

QY 350 tggcaactcactcgtcgttgcgtgagcctcctcctggtccatcgcgtgagcagcctcacc 409  
 Db 1227 CAACCTCACTGGGTGGTGGGACACACTGCTGGCATCGCGTGGAGCGGACCGCAGT 1286

QY 410 tggcgcgaggggggggggggggggggggggggggggggggggggggggggggggg 469  
 Db 1287 TGATGCCGTGAGCTGCACACCGGCTGCCCGTGGCGCGGTGATGCTCATTTGG 1346

QY 470 cggcctggggcggtgcgtcgtcctcctcctcctcctcctcctcctcctcctcctcct 529  
 Db 1347 CGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1406

QY 530 gtgcctcagcgttgcctcactgtcttgcgctcctcctcctcctcctcctcctcctcct 589  
 Db 1407 GTGCCCTGACCGCTCCTCACGATGGACCGCTGCTCAGCGCTCTATTGCGCGCTCT 1466

QY 590 gcgtgcctccttgcgtgggacatcctggcgctctatctgtgcactcagcgcgcactcact 649  
 Db 1467 GGCTCTGTCGAGCGCTGCTTCTTCTGCTCATGTGGTGGTGTGTACACCGCATTTCT 1526

QY 650 gcaggtcagcgcacaaagcgggcgctgcggcgacggcgacggcgacggcgacggcgac 709  
 Db 1527 TCTACCTGGCGGGGAGTGGAGCGCATG-----GCAGAGCATGTCAGCTGCCACCGCG 1581

QY 710 cggacggggcgcgctgcgaagcgcgctgcgtcgtggcctcctcctcctcctcctcctcct 769  
 Db 1582 CTACCGAGA-----GACCAGCTCAGCGTGGTCAAGACTGTGTATCATCC 1628

QY 770 tctcgtccttgcgtatgttgggggggggggggggggggggggggggggggggggggg 829  
 Db 1629 TGGGGCGGTTCGTGTGCTGTGACACAGCGGAGGTGTACTGCTCTGATGTTAG 1688

QY 830 gcccggcgcgacccctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 889  
 Db 1689 GCTGTGAG--TCCGTCAATGTCCTGGCGTTAGAAAGTACTTCTTACTGTTGGCGGAGC 1745

QY 890 ccaactcattctgaaccccatcactacacgctcaccacccgacgactgcgcacgcgc 949  
 Db 1746 CAACCTCACTGTGATGCTGTGTACTTCTTCCGAGATGCTGAGATGCGCGCACCT 1805

QY 950 tctcgcctcgtc 963  
 Db 1806 TCCGCGCGCTTCTC 1819

RESULT 15

US-09-199-737-3  
 ; Sequence 3, Application US/09199737A  
 ; Patent No. 6287788  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bard, Jonathan A.  
 ; APPLICANT: Borowsky, Beth  
 ; APPLICANT: Smith, Kelli E.  
 ; APPLICANT: Branchek, Theresa A.  
 ; APPLICANT: Gerald, Christophe P.G.  
 ; APPLICANT: Jones, Kenneth A.  
 ; TITLE OF INVENTION: DNA Encoding Galanin GALR3 Receptors And Uses Thereof  
 ; FILE REFERENCE: 52241-D-PCT-US  
 ; CURRENT APPLICATION NUMBER: US/09/199,737A  
 ; CURRENT FILING DATE: 1998-11-25  
 ; NUMBER OF SEQ ID NOS: 59  
 ; SOFTWARE: PatentIn Ver. 2.0 - beta  
 ; SEQ ID NO 3  
 ; LENGTH: 1417  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-199-737-3

Query Match 5.68; Score 67.6; DB 4; Length 1417;  
 Best Local Similarity 45.48; Pred. No. 9.2e-05;  
 Matches 244; Conservative 0; Mismatches 294; Indels 0; Gaps 0;

QY 466 gcgcggcgctggggcggtgtcgtcctcctcctcctcctcctcctcctcctcctcctcct 525  
 Db 598 gcgcgagtgggcggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 657

QY 526 ctgggtcctcggagcgttgcctcactgttgcctcactgttgcctcactgttgcctcactgt 595  
 Db 658 tgcggcaccgtgcctacggtgcgtgagctgtgcgtgagctgtgcgtgagctgtgcgtgag 717

QY 586 ttctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 645  
 Db 718 cgcgcgcctcggagcgtggccactcgtcgcgtcgtcgtcgtcgtcgtcgtcgtcgtg 777

QY 646 tactgcaggttacgcgcacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 705  
 Db 778 agcctggcctacggcgacgctgcgttctcgttggcgccggtgggtcccgcgcgcgcg 837

QY 706 acctgcacccggcgcgctgcgaagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 765  
 Db 838 gcggcgcgcgggcgcgcgggcgagggcgagggcgagggcgagggcgagggcgagggcg 897

QY 766 gtgtcctcgtggccttgggcaatgttgggggggggggggggggggggggggggggggg 825  
 Db 898 gcccgcctcctacgcgcgtcgtgtgggtccgcacacgcgcgcgcgcgcgcgcgcgcgc 957

QY 826 gcgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 885  
 Db 958 gggcgcttcgcttcagcccgccacacacacacacacacacacacacacacacacacac 1017

QY 886 atggccaactcacttctgaaccccatcactacacgcgtcaccacacacacacacacacac 945  
 Db 1018 tacgccaactcctgcctcaacccgctcgtacgcgcgtcgcctcgcgcgcacacacacac 1077

QY 946 gcgtcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1003  
 Db 1078 cgcttcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1135

Search completed: January 16, 2002, 23:35:16  
 Job time: 9133 sec



GenCore version 4.5  
 Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2002, 23:38:08 ; Search time 112.43 Seconds  
 (without alignments)  
 517.801 Million cell updates/sec

Title: US-09-842-316-2  
 Perfect score: 2019  
 Sequence: 1 MESGLLRPAVPSEVIVLHYN.....TGSPGAPTAAITLVSEPAAD 398

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues  
 Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 45 summaries

Database : SPTREMBL\_17.\*  
 1: sp-archaea:\*  
 2: sp-bacteria:\*  
 3: sp-fungi:\*  
 4: sp-human:\*  
 5: sp-invertebrate:\*  
 6: sp-mammal:\*  
 7: sp-mhc:\*  
 8: sp-organelle:\*  
 9: sp-phage:\*  
 10: sp-plant:\*  
 11: sp-rodent:\*  
 12: sp-virus:\*  
 13: sp-vertebrate:\*  
 14: sp-unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2019	100.0	398	4	Q9H228
2	1714.5	84.9	400	11	Q9JKM5
3	1698.5	84.1	400	11	Q9QY79
4	1625.5	80.5	387	11	Q9NMN8
5	865.5	42.9	382	11	Q9R235
6	863.5	42.8	382	11	Q9DC35
7	838.5	41.5	382	4	Q9NYN8
8	814.5	40.3	362	13	Q9DDK4
9	814	40.3	326	11	Q9NR55
10	812	40.2	326	11	Q9NR88
11	805	39.9	326	6	Q9BF53
12	804	39.8	326	6	Q9BF72
13	804	39.8	326	11	Q9NR44
14	803	39.8	326	6	Q9BF60
15	803	39.8	326	11	Q9NQ77
16	802	39.7	326	6	Q9BF73
17	802	39.7	326	6	Q9BF46
18	802	39.7	326	11	Q9NQ99
19	801	39.7	326	11	Q9NR11

Q9bf75 myrmecophag  
 Q99nr0 cavia tschu  
 Q9bf79 choleopus h  
 Q9bf78 euphractus  
 Q9bf45 ceratotheri  
 Q9bf44 tapirus ind  
 Q99nr7 muscardinus  
 Q9bf42 leopardus p  
 Q9bf41 panthera on  
 Q9bf76 tamandua te  
 Q9bf74 erinaceus c  
 Q9bf62 cynocephalu  
 Q99nr2 dipodomys h  
 Q9bf67 macroscelid  
 Q99nr3 erethizon d  
 Q9bf69 procavia ca  
 Q9bf65 orycteropus  
 Q99nr9 tamias stri  
 Q9bf70 trichechus  
 Q9bf68 loxodonta a  
 Q9byy4 homo sapien  
 Q9bf57 hylobates c  
 Q9bf77 chaetophrac  
 Q9bf66 elephantulu  
 Q9bf58 macaca mula  
 Q9puq8 fugu rubrip

## ALIGNMENTS

RESULT 1  
 Q9H228  
 ID Q9H228 PRELIMINARY; PRT; 398 AA.  
 AC Q9H228;  
 DT 01-MAR-2001 (Tremblrel. 16, Created)  
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
 DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
 DE SPHINGOSINE 1-PHOSPHATE RECEPTOR EDG-8.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Im D., Lynch K.R.;  
 RT "Characterization of a human sphingosine 1-phosphate receptor, Edg-8";  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF317676; AAG38113.1; -;  
 DR InterPro; IPR000276; GPCR\_Rhodpsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRRHODOPS.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_FL1; UNKNOWN\_1.  
 DR PROSITE; PS0262; G-PROTEIN\_RECEP\_FL2; 1.  
 KW Receptor.  
 SQ SEQUENCE 398 AA; 41775 MW; 34ADC6C40D8250BF CRC64;

Query Match 100.0%; Score 2019; DB 4; Length 398;  
 Best Local Similarity 100.0%; Pred. No. 5.1e-125;  
 Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESGLLRPAVPSEVIVLHYNVTGKRGARYQPGAGLRADAVVCLAVCAFTVLENVAVLV 60  
 |||||||  
 DB 1 MESGLLRPAVPSEVIVLHYNVTGKRGARYQPGAGLRADAVVCLAVCAFTVLENVAVLV 60  
 |||||||  
 QY 61 LGRHPRFAPFLLIGSLTSLDLGAGYAAANTLLSGPLTKLSPALWFAREGGVFVALT 120  
 |||||||  
 DB 61 LGRHPRFAPFLLIGSLTSLDLGAGYAAANTLLSGPLTKLSPALWFAREGGVFVALT 120  
 |||||||  
 QY 121 ASVLSLAIALERSLTWARRGPAPVSRGRTLMAAAWGVSLLLGLPALGVNCLGRD 180  
 |||||||

```

Db 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLIGLLPALGWNCLGRLD 180
QY 181 ACSTVPLVYAKAYVFCVLAFAVGIILAAICALYARIYCOVRANARLPARGCTAGTTSTRA 240
Db 181 ACSTVPLVYAKAYVFCVLAFAVGIILAAICALYARIYCOVRANARLPARGCTAGTTSTRA 240
QY 241 RRKPRSLALLRTLSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 300
Db 241 RRKPRSLALLRTLSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 300
QY 301 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPGSGSQSASAAEASGGRLRCLPPLGDSF 360
Db 301 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPGSGSQSASAAEASGGRLRCLPPLGDSF 360
QY 361 SGERSSPQDGLDTSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 398
Db 361 SGERSSPQDGLDTSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 398

RESULT 2
Q9JKM5 PRELIMINARY; PRT; 400 AA.
ID Q9JKM5 AC Q9JKM5
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SPHINGOSINE 1-PHOSPHATE RECEPTOR EDG-8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20261551; PubMed=10799507;
RA Im D.-S., Heise C.E., Ancellin N., O'Dowd B.F., Shen G.-J.,
RA Heavens R.P., Rigby M.R., Hla T., Mandala S., McAllister G.,
RA George S.R., Lynch K.R.;
RT "Characterization of a novel sphingosine 1-phosphate receptor, Edg-
RT 8."
RL J. Biol. Chem. 275:14281-14286(2000).
DR EMBL; AF233649; AAF35912.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_F1_2; 1.
KW Receptor.
SQ SEQUENCE 400 AA; 42368 MW; 71B398DC4DE0FEB CRC64;

Query Match 84.9%; Score 1714.5; DB 11; Length 400;
Best Local Similarity 87.2%; Pred. No. 4.6e-105;
Matches 346; Conservative 14; Mismatches 34; Indels 3; Gaps 3;

QY 1 MESGLLRPAPVSEVIVLHYNTGKRGARYOPGAGLRADAVVCLAVCAFIENLAVLLV 60
Db 1 MESGLLRPAPVSEVIVLHYNTGKRGARYOPGAGLRADAVVCLAVCAFIENLAVLLV 60
QY 61 LGRHPRFHAPMFLLLGSLTSLDLAAGAAAYANILLSGPLTKLSPALWFAREGGVFALT 120
Db 61 LGRHPRFHAPMFLLLGSLTSLDLAAGAAAYATNILLSGPLTKLSPALWFAREGGVFALT 120
QY 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLIGLLPALGWNCLGRLD 180
Db 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLIGLLPALGWNCLGRLE 180
QY 181 ACSTVPLVYAKAYVFCVLAFAVGIILAAICALYARIYCOVRANARLPARGCTAGTTSTR 239
Db 181 ACSTVPLVYAKAYVFCVLAFAVGIILAAICALYARIYCOVRANARLPARGCTAGTTSTR 239
QY 240 ARKPRSLALLRTLSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 398
Db 240 ARKPRSLALLRTLSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 398
QY 398 SGERSSPQDGLDTSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 398
Db 398 SGERSSPQDGLDTSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 398

```

```

Db 241 SRHTPRSLALLRTLSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 300
QY 300 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPGSGSQSASAAEASGGRLRCLPPLG 357
Db 301 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPGSGSQSASAAEASGGRLRCLPPLG 360
QY 358 SGERSSPQDGLDTSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 394
Db 361 RSSPSEHSCQDGMDSCTSGSPGAATANTLVPD 397

RESULT 3
Q9QY79 PRELIMINARY; PRT; 400 AA.
ID Q9QY79 AC Q9QY79
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE GROWTH FACTOR-REGULATED G PROTEIN-COUPLED RECEPTOR NRG-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99453967; PubMed=10532805;
RA Glickman M., Malek R.L., Wittek-Black A.E., Jacob H.J., Lee N.H.;
RT "Molecular cloning, tissue-specific expression, and chromosomal
RT localization of a novel nerve growth factor-regulated G-protein-
RT coupled receptor, nrg-1."
RL Mol. Cell. Neurosci. 14:141-152(1999).
DR EMBL; AF115249; AAF15395.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_F1_2; 1.
KW Receptor.
SQ SEQUENCE 400 AA; 42368 MW; 72B46CCE3BD24D4B CRC64;

Query Match 84.1%; Score 1698.5; DB 11; Length 400;
Best Local Similarity 86.1%; Pred. No. 5.1e-104;
Matches 342; Conservative 16; Mismatches 36; Indels 3; Gaps 3;

QY 1 MESGLLRPAPVSEVIVLHYNTGKRGARYOPGAGLRADAVVCLAVCAFIENLAVLLV 60
Db 1 MESGLLRPAPVSEVIVLHYNTGKRGARYOPGAGLRADAVVCLAVCAFIENLAVLLV 60
QY 61 LGRHPRFHAPMFLLLGSLTSLDLAAGAAAYANILLSGPLTKLSPALWFAREGGVFALT 120
Db 61 LGRHPRFHAPMFLLLGSLTSLDLAAGAAAYATNILLSGPLTKLSPALWFAREGGVFALT 120
QY 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLIGLLPALGWNCLGRLD 180
Db 121 ASVLSLLAIALERSLTWARGPAPVSSRGRTLAMAAAANGVSLIGLLPALGWNCLGRLE 180
QY 181 ACSTVPLVYAKAYVFCVLAFAVGIILAAICALYARIYCOVRANARLPARGCTAGTTSTR 239
Db 181 ACSTVPLVYAKAYVFCVLAFAVGIILAAICALYARIYCOVRANARLPARGCTAGTTSTR 239
QY 240 ARKPRSLALLRTLSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 398
Db 240 ARKPRSLALLRTLSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 398
QY 398 SGERSSPQDGLDTSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 398
Db 398 SGERSSPQDGLDTSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANS 398

```

```

RESULT 4
Q99MN8
ID Q99MN8 PRELIMINARY; PRT; 387 AA.
AC Q99MN8;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE SPHINGOSINE-1-PHOSPHATE RECEPTOR LPB4 (FRAGMENT).
GN LPB4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS WEBSTER/NIH;
RA Yang A.H., Zhang G., Chun J.J.M.;
RT "Molecular cloning of the mouse sphingosine-1-phosphate receptor gene,
RT Lpb4."
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327535; AAK15485.1; -.
KW NON_TER
FT
SQ SEQUENCE 387 AA; 40953 MW; B56B47672BF5E977 CRC64;

Query Match 80.5%; Score 1625.5; DB 11; Length 387;
Best Local Similarity 86.1%; Pred. No. 3e-99;
Matches 329; Conservative 14; Mismatches 36; Indels 3; Gaps 3;

QY 14 VIVLHNTGKLGARYQPGAGLRADAVVCLAVCAFIIVLENLAVLVLRHPRFHAPMFL 73
Db 1 VIVLHNTGKLGARYQPGAGLRADAVVCLAVCAFIIVLENLAVLVLRHPRFHAPMFL 60

QY 74 LLGSLTSLDLGAGAAATNLLSGPLTKLSPALWFAREGGVFVALTASVLSLAIALER 133
Db 61 LLGSLTSLDLGAGAAATNLLSGPLTKLSPALWFAREGGVFVALTASVLSLAIALER 120

QY 134 SLTMARGPAPVSSRRTLMAAAAGVSLGLLPGALWNCGLRDLACSTVLPYAKAY 193
Db 121 HLTMARGPAPASRARTLMAAAAGVSLGLLPGALWNCGLRDLACSTVLPYAKAY 180

QY 194 VLFVCLAFVILAAICALYARIYCOVRANARLPPAGT-AGTTTRARRRPSRLALLRT 252
Db 181 VLFVCLAFVILAAICALYARIYCOVRANARLPPAGT-AGTTTRARRRPSRLALLRT 240

QY 253 LSVVLLAFVACMGPLFLLLLDVPACPARTCPVLLQADPFLGLAMANSLLNPIIYFTNDR 312
Db 241 LSVVLLAFVACMGPLFLLLLDVPACPARTCPVLLQADPFLGLAMANSLLNPIIYFTNDR 300

QY 313 LRHALLRLVCCGRHSCGRDPSCS-QQSASAAEAS-GGLRRLCPGLDGSFSGSERSSPQR 370
Db 301 LRHALLRLVCCGRHSCGRDPSCS-QQSASAAEAS-GGLRRLCPGLDGSFSGSERSSPQR 360

QY 371 DGLDTSGTSGCAPTAARTLV 392
Db 361 DGLDTSGTSGCAPTAARTLV 382

RESULT 5
Q9R235
ID Q9R235 PRELIMINARY; PRT; 382 AA.
AC Q9R235;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE LYOPHOSPHOLIPID RECEPTOR BL.
GN EDGI OR LPB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN
RP SEQUENCE FROM N.A.
RC STRAIN=125SVJ;
RX MEDLINE=9513220; PubMed=9931453;
RT Zhang G., Contos J.J.A., Weiner J.A., Fukushima N., Chun J.;
RT "Comparative analysis of three murine G-protein coupled receptors
RT activated by sphingosine-1-phosphate."
RL Gene 227:89-99(1999)
DR EMBL; AF108019; AAD16975.1; -.
DR MGI; MGI:1096355; Edgi.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 382 AA; 42613 MW; 2EB4B974E9FBF39C CRC64;

Query Match 42.9%; Score 865.5; DB 11; Length 382;
Best Local Similarity 48.5%; Pred. No. 1.8e-49;
Matches 101; Conservative 66; Mismatches 105; Indels 21; Gaps 6;

QY 13 EVIVHNTGKLR-GARYQPGAGLRADAVVCLAVCAFIIVLENLAVLVLRHPRFHAPM 71
Db 23 DIIVRHNTGKLNIGA--EKDHGKLTSTVVFILCCFIILNIFVLLTIWTKKFRPM 80

QY 72 ELLGSLTSLDLGAGAAATNLLSGPLTKLSPALWFAREGGVFVALTASVLSLAIAL 131
Db 81 YFIGNLALSDLLAGVAYTANLLSGATTYKLTTPAQWFLREGSMFVALSVFSLAIAI 140

QY 132 ERLTMARGPAPVSSRRTLMAAAAGVSLGLLPGALWNCGLRDLACSTVLPYAK 191
Db 141 ERYTMKLMKLNHNGSSRSFLLISACVWISLGLPSMGWNCISSLSGCTVLPYHK 200

QY 192 AYVLCVAFVILAAICALYARIYCOVRANARLPPAGT-AGTTTRARRRPSRLALLR 251
Db 201 HYLFCTVFTVLLLSIVILYCRYISLVTRSRRLTFRK----NISKASRSSEKSLALK 256

QY 252 TLSVLLAFVACMGPLFLLLLDVPACPARTCPVLLQADPFLGLAMANSLLNPIIYFTN 311
Db 257 TVIVLSVFIACWAPLFILLLDVCKAKTCDILYKAEYFLVLAIVNSGPNPIIYFTNK 316

QY 312 DURHALLRLVCCGRHSCGRDPSCSQAASAEAGGLRRLCPGLDGSFSGSERS-PQR 370
Db 317 EMRRAPFIRIV-----SCCKCPNG-----DSAGKFKRPIIPGMEFSRSKSDNSHPQK 363

QY 371 DGLDTSGTSGSPG 383
Db 364 DGDNDPFTINSSG 376

RESULT 6
Q9DC35
ID Q9DC35 PRELIMINARY; PRT; 382 AA.
AC Q9DC35;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE ENDOTHELIAL DIFFERENTIATION SPHINGOLIPID G-PROTEIN-COUPLED RECEPTOR 1.
GN EDGI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=LUNG;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

```





```

DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
KW Receptor. 362 AA; 40583 MW; 43A398871DA2D9D4 CRC64;
SQ SEQUENCE 362 AA; 40583 MW; 43A398871DA2D9D4 CRC64;

Query Match 40.3%; Score 814.5; DB 13; Length 362;
Best Local Similarity 47.1%; Pred. No. 3.8e-46;
Matches 172; Conservative 59; Mismatches 107; Indels 27; Gaps 7;

QY 11 VSEVIVHNYTGKRGARYQDAGLRADAVVCLAVCAFIIVLENLAVLLVGRPHRAP 70
DB 1 MDLIARHNYFTGKRVKHDP--GLKADSVYFIIVCCFIILENVLLTWTTRTKFHKP 58
QY 71 MFLLGSLTSLDLAGAAYANILLSPGLTKLSPALWFAREGGVVFALTASVLSLAI 130
DB 59 MYFIGNLALSLLAGVYVYANILLSGANTYKLTPTQFFREGSMFVALAASVLSLAI 118
QY 131 LERSLTMRARGPAPVSSRGRT---LMAAAAGVSLLLGLLPALGWNCGLRDLACSTVLP 187
DB 119 IEHLTMLK---MKLHNGKTCRVFMLISTVNFIAAILGGLPVMGNCIDSIINNCSTVLP 175
QY 188 LYAKAVYLCVLAFAVGIILAAICARYICOVANARLRLPARPGTAGTSTRARKPRSL 247
DB 176 LYHKAVILECTTVFVILMAVILVARIYALVTRSRKLVKRVKANGSKSEK--SM 233
QY 248 ALLRTLSVLLAFVACWGPFLFLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPIYT 307
DB 234 ALLKTVIIVLSFCIACWAPFLFLLLLDVACOTLCSILYKAWEFLALAVLANSANPLIYT 293
QY 308 LTNRLDLRHALLVCCGRHSCGRDPSGQASAAEAGGLRCLPGLDGSFSGSERS 367
DB 294 LRSNEMRRAFIAKMLNCG-----VCVPSGKFSRPI-MGAEFSSKSDNSS 337
QY 368 -PORD 371
DB 338 HPNKD 342

RESULT 9
Q99NR5 PRELIMINARY; PRT; 326 AA.
AC Q99NR5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE EDG1 (FRAGMENT).
GN EDG1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011706; AAK01975.1; -.
FT NON_TER 1
FT NON_TER 326
SQ SEQUENCE 326 AA; 36574 MW; 7B6DF34BD398B1F7 CRC64;

Query Match 40.28; Score 812; DB 11; Length 326;
Best Local Similarity 49.1%; Pred. No. 5.1e-46;
Matches 166; Conservative 61; Mismatches 93; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLLVGRPHRAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GIKLTSVVFILICCFIILENFVLLTWTKTKFHRPMYFIGNLALSDLLAGVAYTANLL 63
QY 95 LSGPLTKLSPALWFAREGGVVFALTASVLSLAIALERSLTMRARGPAPVSSRGRTLAM 154
DB 64 LSGATTYKLTQAQFREGSMFVALSASFSLAIATERYITLTKMLHNSNSRSRFL 123
QY 155 AAAAGVSLLLGLLPALGWNCGLRDLACSTVLPYAKAVYLCVLAFAVGIILAAIC 214
DB 124 ISACWVSLIILGGLPIMGWNCIGTLSSCSIVLPYHKHYILFCTVFTLLLSIVILYCR 183
QY 215 IYQVRAWARLRLPARPGTAGTSTRARKPRSLALRTLSVLLAFVACWGPFLFLLLLD 274
DB 184 IYSLVTRSRRLTRK----NISKASRSSEKSLALLKTVIIVLSVFIACWAPFLFLL 239
QY 275 VACPARTCPVLLQADPFLGLAMANSLLNPIYTITNDRHALLRLVCCGRHSCGRDPSG 334
DB 240 VGCKAKTCDILYKAIEYFLVAVLNSGNTPIYTITNEMRRAPRIIV-----SCKCPNG 294
QY 335 SQQSASAAEAGGLRCLPGLDGSFSGSERS--PORD 371
DB 295 -----DSAGKFKRPIPGMEFSSKSDNSSHPQKD 324

RESULT 10
Q99NR8 PRELIMINARY; PRT; 326 AA.
AC Q99NR8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE EDG1 (FRAGMENT).
GN EDG1.
OS Castor canadensis (American beaver).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Castoridae; Castor.
OX NCBI_TaxID=51338;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011703; AAK01972.1; -.
FT NON_TER 1
FT NON_TER 326
SQ SEQUENCE 326 AA; 36552 MW; 4BA3EEC6266D1AFA CRC64;

Query Match 40.28; Score 812; DB 11; Length 326;
Best Local Similarity 49.1%; Pred. No. 5.1e-46;
Matches 166; Conservative 61; Mismatches 93; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLLVGRPHRAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GIKLTSVVFILICCFIILENFVLLTWTKTKFHRPMYFIGNLALSDLLAGVAYTANLL 63
QY 95 LSGPLTKLSPALWFAREGGVVFALTASVLSLAIALERSLTMRARGPAPVSSRGRTLAM 154
DB 64 LSGATTYKLTQAQFREGSMFVALSASFSLAIATERYITLTKMLHNSNSRSRFL 123
QY 155 AAAAGVSLLLGLLPALGWNCGLRDLACSTVLPYAKAVYLCVLAFAVGIILAAIC 214
DB 124 ISACWVSLIILGGLPIMGWNCIGTLSSCSIVLPYHKHYILFCTVFTLLLSIVILYCR 183
QY 215 IYQVRAWARLRLPARPGTAGTSTRARKPRSLALRTLSVLLAFVACWGPFLFLLLLD 274
DB 184 IYSLVTRSRRLTRK----NISKASRSSEKSLALLKTVIIVLSVFIACWAPFLFLL 239
QY 275 VACPARTCPVLLQADPFLGLAMANSLLNPIYTITNDRHALLRLVCCGRHSCGRDPSG 334
DB 240 VGCKAKTCDILYKAIEYFLVAVLNSGNTPIYTITNEMRRAPRIIV-----SCKCPNG 294
QY 335 SQQSASAAEAGGLRCLPGLDGSFSGSERS--PORD 371
DB 295 -----DSGKFKRPIPGMEFSSKSDNSSHPQKD 324

```

```

RESULT 11
Q9BF63 ID Q9BF63 PRELIMINARY; PRT; 326 AA.
AC Q9BF63;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE EDG1 (FRAGMENT).
GN EDG1.
OS Ochocona hyperborea (northern pika).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Ochotonidae; Ochotona.
OX NCBI_TaxID=130834;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011717; AAK01985.1; -
FT NON_TER 1 326
FT NON_TER 326
SQ SEQUENCE 326 AA; 36468 MW; 2DEC66EF94397F97 CRC64;

Query Match 39.9%; Score 805; DB 6; Length 326;
Best Local Similarity 48.8%; Pred. No. 1.5e-45;
Matches 165; Conservative 62; Mismatches 93; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLVGRHPRFHAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GKRTSVVFILCCFIILENIFVLLTIWTKKFRHPRMYFIGNLSDLLAGVAYIANLL 63

QY 95 LSGPLTKLSPALWFAREGGVFVALTSVLSLALAIERSITMARRGPAPVSSRGRTLAM 154
DB 64 LSGATTKLTPAQWFLREGSMFALSASFVLSLLAIAIERITMLKMLHNGSNSFRSFL 123

QY 155 AAAGVSVLLGLLPALGNCLGRDACSTVLPYAKAYVLCVAFVGLAICAALYAR 214
DB 124 ISACWVTSILGLPIMGWNCMTLSLSCSTVLPYKHRYILFCTVFTLLLSIVILYCR 183

QY 215 IYCOVRANRRLLPARPGTAGTSTRARRKPRSLALLTSLVLLAFVACWGPLFLLLLD 274
DB 184 IYSLVTRSRRLTFRK---NASKASRSSEKSLALLTVIIVLSVFIACWAPFILLLLD 239

QY 275 VACPARTCVLLQADPFGLAMANSLLNPIIYTLNRDLRALRLVCCGRHSCGRDPSG 334
DB 240 VGCKVKTCILFKAIEYFLVAVLNSGTNPYIITLTNEMERRAFIRIN-----SCCKCPSG 294

QY 335 SQSASAAEASGGLRCLPGLDGSFSGSERSS--PQRD 371
DB 295 -----DSAGKVRPIIAGVFEFSRKSNDSSHPQRD 324

Query Match 39.9%; Score 805; DB 6; Length 326;
Best Local Similarity 48.8%; Pred. No. 1.5e-45;
Matches 165; Conservative 62; Mismatches 93; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLVGRHPRFHAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GKRTSVVFILCCFIILENIFVLLTIWTKKFRHPRMYFIGNLSDLLAGVAYIANLL 63

QY 95 LSGPLTKLSPALWFAREGGVFVALTSVLSLALAIERSITMARRGPAPVSSRGRTLAM 154
DB 64 LSGATTKLTPAQWFLREGSMFALSASFVLSLLAIAIERITMLKMLHNGSNSFRSFL 123

QY 155 AAAGVSVLLGLLPALGNCLGRDACSTVLPYAKAYVLCVAFVGLAICAALYAR 214
DB 124 ISACWVTSILGLPIMGWNCMTLSLSCSTVLPYKHRYILFCTVFTLLLSIVILYCR 183

QY 215 IYCOVRANRRLLPARPGTAGTSTRARRKPRSLALLTSLVLLAFVACWGPLFLLLLD 274
DB 184 IYSLVTRSRRLTFRK---NASKASRSSEKSLALLTVIIVLSVFIACWAPFILLLLD 239

QY 275 VACPARTCVLLQADPFGLAMANSLLNPIIYTLNRDLRALRLVCCGRHSCGRDPSG 334
DB 240 VGCKVKTCILFKAIEYFLVAVLNSGTNPYIITLTNEMERRAFIRIN-----SCCKCPSG 294

QY 335 SQSASAAEASGGLRCLPGLDGSFSGSERSS--PQRD 371
DB 295 -----DSAGKVRPIIAGVFEFSRKSNDSSHPQRD 324

RESULT 12
Q9BF72 ID Q9BF72 PRELIMINARY; PRT; 326 AA.
AC Q9BF72;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE EDG1 (FRAGMENT).
GN EDG1.
OS Sorex araneus (Eurasian common shrew) (European shrew).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Insectivora; Soricidae; Sorex.
OX NCBI_TaxID=42254;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;

```

```

RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011694; AAK01963.1; -
FT NON_TER 1 326
FT NON_TER 326
SQ SEQUENCE 326 AA; 36577 MW; 5BBB6812E3F635B7 CRC64;

Query Match 39.8%; Score 804; DB 6; Length 326;
Best Local Similarity 48.8%; Pred. No. 1.7e-45;
Matches 165; Conservative 60; Mismatches 95; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLVGRHPRFHAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GKRTSVVFILCCFIILENIFVLLTIWTKKFRHPRMYFIGNLSDLLAGVAYIANLL 63

QY 95 LSGPLTKLSPALWFAREGGVFVALTSVLSLALAIERSITMARRGPAPVSSRGRTLAM 154
DB 64 LSGATTKLTPAQWFLREGSMFALSASFVLSLLAIAIERITMLKMLHNGSNSFRSFL 123

QY 155 AAAGVSVLLGLLPALGNCLGRDACSTVLPYAKAYVLCVAFVGLAICAALYAR 214
DB 124 ISACWVTSILGLPIMGWNCMTLSLSCSTVLPYKHRYILFCTVFTLLLSIVILYCR 183

QY 215 IYCOVRANRRLLPARPGTAGTSTRARRKPRSLALLTSLVLLAFVACWGPLFLLLLD 274
DB 184 IYSLVTRSRRLTFRK---NASKASRSSEKSLALLTVIIVLSVFIACWAPFILLLLD 239

QY 275 VACPARTCVLLQADPFGLAMANSLLNPIIYTLNRDLRALRLVCCGRHSCGRDPSG 334
DB 240 VGCKVKTCILFKAIEYFLVAVLNSGTNPYIITLTNEMERRAFIRIN-----SCCKCPSG 294

QY 335 SQSASAAEASGGLRCLPGLDGSFSGSERSS--PQRD 371
DB 295 -----DSAGKVRPIIAGVFEFSRKSNDSSHPQRD 324

RESULT 13
Q99NR4 ID Q99NR4 PRELIMINARY; PRT; 326 AA.
AC Q99NR4;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE EDG1 (FRAGMENT).
GN EDG1.
OS Hystrix brachyurus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Hystricidae; Hystrix.
OX NCBI_TaxID=143286;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011708; AAK01976.1; -
FT NON_TER 1 326
FT NON_TER 326
SQ SEQUENCE 326 AA; 36546 MW; 3E66CA765F725BA9 CRC64;

Query Match 39.8%; Score 804; DB 11; Length 326;
Best Local Similarity 48.8%; Pred. No. 1.7e-45;
Matches 165; Conservative 60; Mismatches 95; Indels 18; Gaps 4;

QY 35 GLRADAVVCLAVCAFIIVLENLAVLVGRHPRFHAPFLLGSLTSLDLAGAAYANIL 94
DB 4 GKRTSVVFILCCFIILENIFVLLTIWTKKFRHPRMYFIGNLSDLLAGVAYIANLL 63

```

```

QY 95 LSGPLTLKSPALMFAREGGVVALTSVLSLALALERSLTMARRGPAPVSSRGRTLAM 154
Db 64 LSGATTYKLTPAQWFLREGSMFALSASFSLAIAIERVITLKKMLHNGSNSRSFLL 123
QY 155 AAAAGVSLILGLLPALGNCLGRDACSTVLPYAKAYVLCVLAFAVGIILAAICAYAR 214
Db 124 ISACWVSLILGLLPALGNCLGRDACSTVLPYAKAYVLCVLAFAVGIILAAICAYAR 214
QY 215 IYCOVRANARLPARPAGTGTSTRARRKPSLALLRTLSTVLLAFVACWGPFLFLLLD 274
Db 184 IYSLVTRSRRLTFRK-----SVSKASRSSEKSLALLKTVIIVLSVFIACWAPFLFLLLD 239
QY 275 VACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRDLRHALLRVCCGRHSCGRDPG 334
Db 184 IYSLVTRSRRLTFRK-----SVSKASRSSEKSLALLKTVIIVLSVFIACWAPFLFLLLD 239
QY 240 VGCKVKTCDILFTEYFLVAVLNSGTNPITLTNKMERRAFIRIM-----SCKCPGG 294
QY 335 SQOSASAAEASGRLRCLPPGLDGFSGFSERS-PORD 371
Db 295 -----DSTGKIKRPIIAGVEFSRKSNDSSHPOKD 324

RESULT 14
ID Q9BF60 PRELIMINARY; PRT; 326 AA.
AC Q9BF60;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE EDG1 (FRAGMENT).
GN EDG1.
OS Lemur catta (Ring-tailed lemur).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Strepsirhini; Lemuridae; Lemur.
OX NCBI_TaxID=9447;
RN [1]
RP SEQUENCE FROM N.A..
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011720; AAK01988.1; -.
FT NON_TER 1
FT NON_TER 326
SQ SEQUENCE 326 AA; 36289 MW; B4D31565F3D5A6AD CRC64;

Query Match 39.8%; Score 803; DB 6; Length 326;
Best Local Similarity 49.4%; Pred. No. 2e-45;
Matches 167; Conservative 57; Mismatches 96; Indels 18; Gaps 4;

QY 35 GLRADAVCLAVCAFTVLENLAVLVGRPHRPHAPMFLLSGLSLDLAGAAYANILL 94
Db 4 GLAFTSVFILLICCFIILENIFVLLTWKTKRPHMYFFIGNLALSOLLAGVAYIANLL 63
QY 95 LSGPLTLKSPALMFAREGGVVALTSVLSLALALERSLTMARRGPAPVSSRGRTLAM 154
Db 64 LSGATTYKLTPAQWFLREGSMFALSASFSLAIAIERVITLKKMLHNGSNSRSFLL 123
QY 155 AAAAGVSLILGLLPALGNCLGRDACSTVLPYAKAYVLCVLAFAVGIILAAICAYAR 214
Db 124 ISACWVSLILGLLPALGNCLGRDACSTVLPYAKAYVLCVLAFAVGIILAAICAYAR 214
QY 215 IYCOVRANARLPARPAGTGTSTRARRKPSLALLRTLSTVLLAFVACWGPFLFLLLD 274
Db 184 IYSLVTRSRRLTFRK-----SVSKASRSSEKSLALLKTVIIVLSVFIACWAPFLFLLLD 239
QY 275 VACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRDLRHALLRVCCGRHSCGRDPG 334
Db 240 VGCKVKTCDILFTEYFLVAVLNSGTNPITLTNKMERRAFIRIV-----SCKCPGG 294
QY 335 SQOSASAAEASGRLRCLPPGLDGFSGFSERS-PORD 371
Db 295 -----DSTGKIKRPIIAGVEFSRKSNDSSHPOKD 324

```

```

Db 295 -----DAAGVKRPIVAGVEFSRKSNDSSHPOKD 324

RESULT 15
ID Q99N07 PRELIMINARY; PRT; 326 AA.
AC Q99N07;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE EDG1 (FRAGMENT).
GN EDG1.
OS Agouti taczanowski.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Mystricognathi; Agoutidae; Agouti.
OX NCBI_TaxID=143288;
RN [1]
RP SEQUENCE FROM N.A..
RX MEDLINE=21082082; PubMed=11214319;
RA Murphy W.J., Elzirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,
RA O'Brien S.J.;
RT "Molecular phylogenetics and the origins of placental mammals.";
RL Nature 409:614-618(2001).
DR EMBL; AY011715; AAK01983.1; -.
FT NON_TER 1
FT NON_TER 326
SQ SEQUENCE 326 AA; 36652 MW; 3AD51684BF81A889 CRC64;

Query Match 39.8%; Score 803; DB 11; Length 326;
Best Local Similarity 48.7%; Pred. No. 2e-45;
Matches 164; Conservative 62; Mismatches 93; Indels 18; Gaps 4;

QY 36 LRADAVCLAVCAFTVLENLAVLVGRPHRPHAPMFLLSGLSLDLAGAAYANILL 95
Db 5 VKUTTVFILLICCFIILENIFVLLTWKTKRPHMYFFIGNLALSOLLAGVAYIANLL 64
QY 96 SGLPTLKLSPALMFAREGGVVALTSVLSLALALERSLTMARRGPAPVSSRGRTLAM 155
Db 65 SGATTYKLTPAQWFLREGSMFALSASFSLAIAIERVITLKKMLHNGSNSRSFLLI 124
QY 156 AAAAGVSLILGLLPALGNCLGRDACSTVLPYAKAYVLCVLAFAVGIILAAICAYAR 215
Db 125 SACWVSLILGLLPALGNCLGRDACSTVLPYAKAYVLCVLAFAVGIILAAICAYAR 215
QY 216 YCOVRANARLPARPAGTGTSTRARRKPSLALLRTLSTVLLAFVACWGPFLFLLLDV 275
Db 185 YSLVTRSRRLTFRKNISK-----SRSEKSLALLKTVIIVLSVFIACWAPFLFLLLDV 240
QY 276 ACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRDLRHALLRVCCGRHSCGRDPG 335
Db 241 GCKVKTCDILFTEYFLVAVLNSGTNPITLTNKMERRAFIRIM-----SCKCPGG- 294
QY 336 SQOSASAAEASGRLRCLPPGLDGFSGFSERS-PORD 371
Db 295 -----DSTGKIKRPIIAGVEFSRKSNDSSHPOKD 324

```

Search completed: January 16, 2002, 23:43:27  
Job time: 319 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.  
OM protein - protein search, using sw model  
Run On: January 16, 2002, 23:04:03 ; Search time 58.72 Seconds  
(without alignments)  
152,526 Million cell updates/sec  
Title: US-09-842-316-2  
Perfect score: 2019  
Sequence: 1 MESGLLRPAVPSEVIVLHYN.....TGSPGAPTAARTLVSEPAAD 398  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 212252 seqs, 22503292 residues  
Total number of hits satisfying chosen parameters: 212252  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : Issued\_Patents\_AA.\*  
1: /cgn2.6/ptodata/2/1aa/5A\_COMB.pap.\*  
2: /cgn2.6/ptodata/2/1aa/5B\_COMB.pap.\*  
3: /cgn2.6/ptodata/2/1aa/6A\_COMB.pap.\*  
4: /cgn2.6/ptodata/2/1aa/6B\_COMB.pap.\*  
5: /cgn2.6/ptodata/2/1aa/PCTUS\_COMB.pap.\*  
6: /cgn2.6/ptodata/2/1aa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	853.5	42.3	383	1	US-08-196-989B-4
2	853.5	42.3	383	2	US-08-760-936-4
3	834	41.3	381	2	US-08-845-566-3
4	834	41.3	381	2	US-08-467-948A-28
5	834	41.3	381	3	US-08-852-824-18
6	834	41.3	381	3	US-08-467-947A-28
7	834	41.3	381	5	PCT-US96-10618-4
8	776	38.4	334	1	US-08-118-270-73
9	776	38.4	334	5	PCT-US93-08528-73
10	774.5	38.4	378	4	US-09-082-088-2
11	732.5	36.3	352	1	US-08-196-989B-2
12	732.5	36.3	352	2	US-08-760-936-2
13	631.5	31.3	284	3	US-08-852-824-4
14	548	27.1	509	2	US-08-845-566-1
15	518.5	25.7	364	4	US-08-763-938-2
16	508.5	25.2	351	3	US-08-861-747-2
17	505.5	25.0	393	5	PCT-US96-10618-3
18	502.5	24.9	364	5	PCT-US96-10618-2
19	499.5	24.7	351	3	US-08-789-882-2
20	445.5	22.1	354	4	US-08-325-897-2
21	443	21.9	353	3	US-08-997-803-14
22	339.5	16.8	213	3	US-08-997-803-15
23	339.5	16.8	393	2	US-08-467-948A-4
24	339.5	16.8	393	3	US-08-467-947A-4
25	304.5	15.1	345	1	US-08-118-270-70
26	304.5	15.1	345	5	PCT-US93-08528-70
27	283.5	14.0	325	3	US-08-706-281A-18

28	283.5	14.0	325	4	US-09-097-231-18
29	282.5	14.0	325	1	US-08-671-525B-10
30	282.5	14.0	325	1	US-08-672-109B-10
31	282.5	14.0	325	2	US-08-842-045-10
32	282.5	14.0	325	2	US-08-842-238-10
33	282.5	14.0	325	3	US-08-629-335B-10
34	273	13.5	332	2	US-08-662-560-2
35	273	13.5	332	2	US-08-780-749A-6
36	273	13.5	332	4	US-08-870-511-6
37	273	13.5	332	4	US-08-870-511-12
38	273	13.5	332	4	US-08-870-511-12
39	272	13.5	332	1	US-08-671-525B-8
40	272	13.5	332	1	US-08-672-109B-8
41	272	13.5	332	2	US-08-842-045-8
42	272	13.5	332	2	US-08-842-238-8
43	272	13.5	332	2	US-08-780-749A-2
44	272	13.5	332	3	US-08-629-335B-8
45	272	13.5	332	4	US-08-870-511-2

ALIGNMENTS

RESULT 1  
US-08-196-989B-4  
; Sequence 4, Application US/08196989B  
; Patent No. 5585476  
; GENERAL INFORMATION:  
; APPLICANT: MacLennan, A. John  
; TITLE OF INVENTION: Molecular Cloning and Expression of  
; TITLE OF INVENTION: G-Protein Coupled Receptors  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08196.989B  
; FILING DATE: 15-FEB-1994  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lloyd, Jeff  
; REGISTRATION NUMBER: 35,589  
; REFERENCE/DOCKET NUMBER: MAC-100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 904-375-8100  
; TELEFAX: 904-372-5800  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 383 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-196-989B-4

Query Match 42.3%; Score 853.5; DB 1; Length 383;  
Best Local Similarity 47.6%; Pred. No. 2.5e-58;  
Matches 177; Conservative 66; Mismatches 110; Indels 19; Gaps 5;  
QY 13 EVIVLHYNHTKLRGARYQPGAGLRADAVVCLAVCAFTVLNLAIVLGLRHRFRHAPMF 72  
DB 24 DIIVRHYNHTKGL-NIGVEKDHGKLTAVVFLICCLILLENIFVLLTIWTKKFKRPMY 82  
QY 73 LLGLSLTLDLAGAAYANILLSSPLTLKLSPALWFAREGGVFVALTASVLSLLAIALE 132

Db 83 YFIGNLALSDLLAGVAYTANLLSGATTYKLTTPAQOFLREGSMFVALSASFVSLAIAIE 142  
QY 133 RSLTWARREGPAPVSSRGRTLAMAAAAGVSLILGLLPALGWNCIGRLDACSTVLPYAKA 192  
Db 143 RYITMLKMLKLNHNSRSLFLISACWVLSILGLLPINGWNCISLSSCSTVLPYHKH 202  
QY 193 YVLCFVAFVGIILAAICALYARIYCOVRANARRLPARPCTAGTTSTRARRKPRSLALLRT 252  
Db 203 YILECTVFTLLLSIVLYCRIYSLVTRSRRLTRK----NISKASRSSEKSLALLKT 258  
QY 253 LSVLLAFVACWGLPFLLLLDVACPARCVPVLLQADPGLGLAMANSLLNPIITLTNRD 312  
Db 259 VIIVLSVFIACWAPLFIILLLDVGCKAKTCDILYKAEYFLVLAVALNSGTNPPIIYLTNKE 318  
QY 313 LRHALLRLVCCGRHSGRDPGSGSQSQAASAEASGGLRLCPGLDGSFSGSERS--PORD 371  
Db 319 MRRAFIRII-----SCKKCPNG-----DSAGKFRPIIPGMFEFSKSDNSHSPQKD 365  
QY 372 GLDTSGSTGSPG 383  
Db 366 DGDNPETIMSSG 377

RESULT 2  
US-08-760-936-4  
; Sequence 4, Application US/08760936  
; Patent No. 5856443  
; GENERAL INFORMATION:  
; APPLICANT: MacLennan, A. John  
; TITLE OF INVENTION: Molecular Cloning and Expression of  
; TITLE OF INVENTION: G-Protein Coupled Receptors  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/760,936  
; FILING DATE: December 6, 1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pace, Doran R.  
; REGISTRATION NUMBER: 38,261  
; REFERENCE/DOCKET NUMBER: MAC-100C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 352-373-8100  
; TELEFAX: 352-372-5800  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 383 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-760-936-4

Query Match 42.3%; Score 853.5; DB 2; Length 383;  
Best Local Similarity 47.6%; Pred. No. 2.5e-58;  
Matches 177; Conservative 66; Mismatches 110; Indels 19; Gaps 5;  
QY 13 EVIVLHYNTYKRGARYOPGAGLRADAVVCLAVCAFIIVLENLAVLVIGLRHPRFHAPMF 72  
Db 24 DIIVRHYNTYTKL-NIGVEKDHGKLTSVWFILICLLILENIFVLLTINKTKFHRPHY 82

QY 73 LLIGSTLSDLLAGAAAYANILLSGFLTKLSPALFAREGGVFALTASVLSLAIAIE 132  
Db 83 YFIGNLALSDLLAGVAYTANLLSGATTYKLTTPAQOFLREGSMFVALSASFVSLAIAIE 142  
QY 133 RSLTWARREGPAPVSSRGRTLAMAAAAGVSLILGLLPALGWNCIGRLDACSTVLPYAKA 192  
Db 143 RYITMLKMLKLNHNSRSLFLISACWVLSILGLLPINGWNCISLSSCSTVLPYHKH 202  
QY 193 YVLCFVAFVGIILAAICALYARIYCOVRANARRLPARPCTAGTTSTRARRKPRSLALLRT 252  
Db 203 YILECTVFTLLLSIVLYCRIYSLVTRSRRLTRK----NISKASRSSEKSLALLKT 258  
QY 253 LSVLLAFVACWGLPFLLLLDVACPARCVPVLLQADPGLGLAMANSLLNPIIYLTNRD 312  
Db 259 VIIVLSVFIACWAPLFIILLLDVGCKAKTCDILYKAEYFLVLAVALNSGTNPPIIYLTNKE 318  
QY 313 LRHALLRLVCCGRHSGRDPGSGSQSQAASAEASGGLRLCPGLDGSFSGSERS--PORD 371  
Db 319 MRRAFIRII-----SCKKCPNG-----DSAGKFRPIIPGMFEFSKSDNSHSPQKD 365  
QY 372 GLDTSGSTGSPG 383  
Db 366 DGDNPETIMSSG 377

RESULT 3  
US-08-845-566-3  
; Sequence 3, Application US/08845566  
; Patent No. 5912144  
; GENERAL INFORMATION:  
; APPLICANT: Au-Young, Janice  
; APPLICANT: Guegler, Karl  
; TITLE OF INVENTION: EDG-1 LIKE RECEPTOR  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/845,566  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0271 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 381 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: Genbank  
; CLONE: 181948  
US-08-845-566-3

Query Match 41.3%; Score 834; DB 2; Length 381;  
Best Local Similarity 46.8%; Pred. No. 7.8e-57;

```

Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;
QY 13 EVIVLHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIIVLENLAVLLVLRHPRFHPMF 72
DB 23 DIIVRHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIIVLENLAVLLVLRHPRFHPMF 81
QY 73 LLGSLTSLDLAAGAYANILSGPLTKLSPALWFAREGGVFVALTASVLSLLATALE 132
DB 82 YFIGNLALSLLAGVAYTANLLSGATTYKLTPAQWFLREGSMFVALSASFSLATAIE 141
QY 133 RSLTMARRGPAPVSSRGRTLMAAAAGVSLILGLLPALGWNCGLRDLACSTVLPYAKA 192
DB 142 RYITMKLKHNGSNFRFLISACWVLSILGGPLMGWNCISALSSCSTVLPYHKK 201
QY 193 YVLCVLAFCVILAAICALYARIYCOVRANARLPARTAGTTSTRARRKPRSLALLRT 252
DB 202 YILFCTVFTLLLSIVILYCRISLVTRRSRLTKRNI-----SKASRSSENVALKLT 256
QY 253 LSVVLLAFVACWGPLFLLLLDVACPARTCPVLLQADPFLGLMANSLNPIIYTLNRD 312
DB 257 VIIVLSVFIACHAPFLFLLLDVGCKVKTCDILFRAEYFLVLAVLSNGTNPPIIYTLNKE 316
QY 313 LRHALLRLVCCGRHSCDRPDSGSOASAAEASGGLRCLPPLDGSFSGSERSS-PORD 371
DB 317 MRRAFRIM-----SCKKCPSG-----DSAGKFKRPIIAGMEFSRSKSDNSHPDKD 363
QY 372 GLDTSGSTGSPG 383
DB 364 EGDNPETIMSSG 375

```

```

RESULT 4
US-08-467-948A-28
; Sequence 28, Application US/08467948A
; Patent No. 5998164
; GENERAL INFORMATION:
; APPLICANT: LI, YI
; APPLICANT: CAO, LIANG
; APPLICANT: NI, JIAN
; APPLICANT: GENTZ, REINER
; APPLICANT: BULT, CAROL J.
; APPLICANT: SUTTON III, GRANGER G.
; APPLICANT: ROSEN, CRAIG A.
; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVE., NW, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,948A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04079
; FILING DATE: 30-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: STEFFE, ERIC K.
; REGISTRATION NUMBER: 36,688
; REFERENCE/DOCKET NUMBER: 1488.1140003/EKS/KLM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-2500
; TELEFAX: 202-371-2540
; INFORMATION FOR SEQ ID NO: 28:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-467-948A-28

Query Match 41.3%; Score 834; DB 2; Length 381;
Best Local Similarity 46.8%; Pred. No. 7.8e-57;
Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;
QY 13 EVIVLHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIIVLENLAVLLVLRHPRFHPMF 72
DB 23 DIIVRHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIIVLENLAVLLVLRHPRFHPMF 81
QY 73 LLGSLTSLDLAAGAYANILSGPLTKLSPALWFAREGGVFVALTASVLSLLATALE 132
DB 82 YFIGNLALSLLAGVAYTANLLSGATTYKLTPAQWFLREGSMFVALSASFSLATAIE 141
QY 133 RSLTMARRGPAPVSSRGRTLMAAAAGVSLILGLLPALGWNCGLRDLACSTVLPYAKA 192
DB 142 RYITMKLKHNGSNFRFLISACWVLSILGGPLMGWNCISALSSCSTVLPYHKK 201
QY 193 YVLCVLAFCVILAAICALYARIYCOVRANARLPARTAGTTSTRARRKPRSLALLRT 252
DB 202 YILFCTVFTLLLSIVILYCRISLVTRRSRLTKRNI-----SKASRSSENVALKLT 256
QY 253 LSVVLLAFVACWGPLFLLLLDVACPARTCPVLLQADPFLGLMANSLNPIIYTLNRD 312
DB 257 VIIVLSVFIACHAPFLFLLLDVGCKVKTCDILFRAEYFLVLAVLSNGTNPPIIYTLNKE 316
QY 313 LRHALLRLVCCGRHSCDRPDSGSOASAAEASGGLRCLPPLDGSFSGSERSS-PORD 371
DB 317 MRRAFRIM-----SCKKCPSG-----DSAGKFKRPIIAGMEFSRSKSDNSHPDKD 363
QY 372 GLDTSGSTGSPG 383
DB 364 EGDNPETIMSSG 375

```

```

RESULT 5
US-08-852-824-18
; Sequence 18, Application US/08852824C
; Patent No. 6060272
; GENERAL INFORMATION:
; APPLICANT: LI, et al.
; TITLE OF INVENTION: Human G-Protein Coupled Receptors
; FILE REFERENCE: 1488.1220000
; CURRENT APPLICATION NUMBER: US/08/852,824C
; CURRENT FILING DATE: 1997-05-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 18
; LENGTH: 381
; TYPE: PRT
; ORGANISM: genomic
; US-08-852-824-18

Query Match 41.3%; Score 834; DB 3; Length 381;
Best Local Similarity 46.8%; Pred. No. 7.8e-57;
Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;
QY 13 EVIVLHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIIVLENLAVLLVLRHPRFHPMF 72
DB 23 DIIVRHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIIVLENLAVLLVLRHPRFHPMF 81
QY 73 LLGSLTSLDLAAGAYANILSGPLTKLSPALWFAREGGVFVALTASVLSLLATALE 132
DB 82 YFIGNLALSLLAGVAYTANLLSGATTYKLTPAQWFLREGSMFVALSASFSLATAIE 141

```



QY 133 RSLTWARGPAPVSSRGRTLAMAAAGVSVLLGLLPLAGWNCGLRDLACSTVLPYAKA 192  
 Db 142 RYITMLKMLKUNGNNFLLISACWVSLILGGLPIMGWNCISALSSCSTVLPYHKG 201  
 QY 193 YVLCVLAFAVGLAAICALYARIYCOVRANARLPAPGTAGTSTRARRKPRSLALLRT 252  
 Db 202 YILFCTVFTLLLSIVILYCRYISLVTRSRRLTKNI-----SKASRSSENVALLKT 256  
 QY 253 LSVLLAFVACWGLPFLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRD 312  
 Db 257 VIIVLSVFIACWAPFLFLLLDVGCWKTCIDLFRAYEFLVLAVALNSGTNPPIIYTLNKE 316  
 QY 313 LRHALLRVCGRHSGRDPGSGSQSASAAEAGGLRCLPGLDGSFSGSERSS-PORD 371  
 Db 317 MRRAFIRM-----SCCKCPSG-----DSAGKFRPIIAGMEFSRKSNDSSHPOKD 363  
 QY 372 GLDTSSTGSGPG 383  
 Db 364 EGDNPETIMSSG 375

RESULT 6

; Sequence 28, Application US/08467947A  
 ; Patent No. 6090575  
 ; GENERAL INFORMATION:  
 ; APPLICANT: LI, YI  
 ; APPLICANT: CAO, LIANG  
 ; APPLICANT: NI, JIAN  
 ; APPLICANT: GENTZ, REINER  
 ; APPLICANT: BULT, CAROL J.  
 ; APPLICANT: SUTTON III, GRANGER G.  
 ; APPLICANT: ROSEN, CRAIG A.  
 ; TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein  
 ; NUMBER OF SEQUENCES: 30  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.  
 ; STREET: 1100 NEW YORK AVE., NW, SUITE 600  
 ; CITY: WASHINGTON  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: FLOPPY DISK  
 ; COMPUTER: IBM PC COMPATIBLE  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PATENTIN RELEASE #1.0, VERSION #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/467,947A  
 ; FILING DATE: 06-JUN-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/04079  
 ; FILING DATE: 30-MAR-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: STEFFE, ERIC K.  
 ; REGISTRATION NUMBER: 36,688  
 ; REFERENCE/DOCKET NUMBER: 1488.1140002/EKS/KLM  
 ; TELEPHONE: 202-371-2600  
 ; TELEFAX: 202-371-2540  
 ; INFORMATION FOR SEQ ID NO: 28:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 381 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: not relevant  
 ; MOLECULE TYPE: peptide  
 ; US-08-467-947A-28

Query Match 41.3%; Score 834; DB 3; Length 381;

Best Local Similarity 46.8%; Pred. No. 7.8e-57;  
 Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;  
 QY 13 EVIIVHNTGKLGARYOPGAGLRADAVVCLAVCAFIYVLENLAVLLVLRHPRHAPMF 72  
 Db 23 DIIVRYHNTGKUNISADKENS-IKLTSVFILICCFILENFVLLTWTKKPHRPMY 81  
 QY 73 LLGSLTSLDLLAGAAAYANILLSGPLTLKSPALWFAREGGVFVALTASVLSLAI 132  
 Db 82 YFIGNLALSDLLAGVAYTANLLSGATTYKLTTPAQWFLREGSMFVALSASFLLAIAIE 141  
 QY 133 RSLTWARGPAPVSSRGRTLAMAAAGVSVLLGLLPLAGWNCGLRDLACSTVLPYAKA 192  
 Db 142 RYITMLKMLKUNGNNFLLISACWVSLILGGLPIMGWNCISALSSCSTVLPYHKG 201  
 QY 193 YVLCVLAFAVGLAAICALYARIYCOVRANARLPAPGTAGTSTRARRKPRSLALLRT 252  
 Db 202 YILFCTVFTLLLSIVILYCRYISLVTRSRRLTKNI-----SKASRSSENVALLKT 256  
 QY 253 LSVLLAFVACWGLPFLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRD 312  
 Db 257 VIIVLSVFIACWAPFLFLLLDVGCWKTCIDLFRAYEFLVLAVALNSGTNPPIIYTLNKE 316  
 QY 313 LRHALLRVCGRHSGRDPGSGSQSASAAEAGGLRCLPGLDGSFSGSERSS-PORD 371  
 Db 317 MRRAFIRM-----SCCKCPSG-----DSAGKFRPIIAGMEFSRKSNDSSHPOKD 363  
 QY 372 GLDTSSTGSGPG 383  
 Db 364 EGDNPETIMSSG 375

RESULT 7

PCT-US96-10618-4  
 ; Sequence 4, Application PC/TUS9610618  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Coleman, Roger  
 ; APPLICANT: Guegler, Karl J.  
 ; APPLICANT: Au-Yang, Janice  
 ; APPLICANT: Bandman, Olga  
 ; APPLICANT: Seilhamer, Jeffrey J.  
 ; TITLE OF INVENTION: A NOVEL HUMAN EDG-2 RECEPTOR HOMOLOG  
 ; NUMBER OF SEQUENCES: 6  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
 ; STREET: 3174 Porter Drive  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94304  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ Version 1.5  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US96/10618  
 ; FILING DATE: 20-JUN-1996  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 60/000,352  
 ; FILING DATE: 20-JUN-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/567,817  
 ; FILING DATE: 06-DEC-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Glaister, Debra J.  
 ; REGISTRATION NUMBER: 33,888  
 ; REFERENCE/DOCKET NUMBER: PF-0042 PCT  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-855-0555  
 ; TELEFAX: 415-845-4166  
 ; TELEX:

```

; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 119130
; PCT-US96-10618-4

Query Match 41.3%; Score 834; DB 5; Length 381;
Best Local Similarity 46.8%; Pred. No. 7.8e-57;
Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;

QY 13 EVVLHNYTGKRGARVOPGAGLRADAVVCLAVCAFTVLENLAVLLVGLGHRPHAPMF 72
DB 23 DIIVRHNYTGKLNISADKENS-IKLTSVFILICCFIILENIFVLTITWTKKFRPMY 81
QY 73 LLGSLTSLDLGAGYAANILLSPGLKSLPALWFAREGGVFVALTSVLSLALALE 132
DB 82 YFIGNLSDLLAGVAYTANLLSGATTYKLTQAQWFLREGSMFVALSASVFSLLAIE 141
QY 133 RSLTHARRGPAPVSRGRTLAMAAAGVSLLLGLLPALGNCGLRLDACSTVLPYAKA 192
DB 142 RYITLMLKWLHNGSNFRLFLISACWVLSILGGLPIMGWNCISALSSGCTVLPYHKH 201
QY 193 YVLCVAVFGVLAALCAIYARIYQVRRANRRPAPGTAGTSTTRARRKPSLALLRT 252
DB 202 YILFCTVFTLLLSIVLYCRISLVTRSRRLTFKNT-----SKASRSSENVALLKT 256
QY 253 LSVLLAFVACGPFLLLLLDVACPARTCPVLLQADPFGLAMANSLLNPIIYLTNRD 312
DB 257 VVILSVFIAWAPLFIILLDVGCKVKTCDFLFAEYFLVLAIVNSGTNPPIYTLTKE 316
QY 313 LRHALLRLVCCGRHSCGRDPSSQOSASAEASGGLRCLPPLGDSFGSERS-PORD 371
DB 317 MRRAFIRM-----SCKCPG-----DSAGKFRPIIAGMEFSRSKSDNSHPQKD 363
QY 372 GLDTSGTSGSPG 383
DB 364 EGDNPETIMSSG 375

RESULT 8
US-08-118-270-73
; Sequence 73, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US, 07/943,236

```

```

; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 73:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 334 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-118-270-73

Query Match 38.4%; Score 776; DB 1; Length 334;
Best Local Similarity 46.4%; Pred. No. 1.9e-52;
Matches 162; Conservative 64; Mismatches 95; Indels 28; Gaps 7;

QY 40 AVVCLAVCAFTVLENLAVLLVGLGHRPHAPMFLLGLSLTSLDLGAGYAANILLSGPL 99
DB 3 SVYFLICCFIILENIFVLTITWTKKFRPMYFIIGNLSLIALSGLIAGVAYTANLLSGAT 62
QY 100 TLKLSPALWFAREGGVFVALTSVLSLALALERSLTMAR---RGPAPVSSRCRTLAMAA 156
DB 63 TYKLPAPQWFLREGSMFVALSLSVFSLLAIAIERIYITMLKMLHG-----SNNRFLFLIS 118
QY 157 AAWGVSLLLGLLPALGNCGLRLDACSTVLPYAKAYVLCVLA-FVGLIAAICALYARI 215
DB 119 ACWVLSILGGLPIMGWNCISALSSGCTVLPYHKHYLFTLIVFTLLLSIVLYCRI 178
QY 216 YCOVRRANRRPAPGTAGTSTTRARRKPSLALLRTLSVLLAFVACGPFLLLLLDV 275
DB 179 YSLVTRSRRLTFKNT-----SKASRSSENVALLKTIVILSVFIAWAPLFIILLDV 233
QY 276 ACAPATCPVLLQADPFGLAMANSLLNPIIYLTNRDLRHALLRLVCCGRHSCGRDPSSG 335
DB 234 GCKVTKCDILFAEYFLVLAIVNSGTNPPIIYLTNEMRRAFIRIMCC-----KCPG- 286
QY 336 QOSASAEASGGLRCLPPLGDSFGSERS-PORDGLDTSGTSGSPG 383
DB 287 -----DSAGKFRPIIAGMEFSRSKSDNSHPQKDEGDNPETIMSSG 328

RESULT 9
PCT-US93-08528-73
; Sequence 73, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236

```

; FILING DATE: 10-SEP-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Townsend, Kevin G.  
; REGISTRATION NUMBER: 34,033  
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-628-5197  
; TELEFAX: 202-737-3528  
; TELEX: 248633  
; INFORMATION FOR SEQ ID NO: 73:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 334 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
PCT-US93-08528-73

Query Match 38.4%; Score 776; DB 5; Length 334;  
Best Local Similarity 46.4%; Pred. No. 1.9e-52;  
Matches 162; Conservative 64; Mismatches 95; Indels 28; Gaps 7;  
QY 40 AVYCAVCAFTVLENLAVLLVGLRPHAPMFLILGSLTSLDLAGAAYAAANILLGSL 99  
DB 3 SVFLLICCFILLENIFVLLTIKTKKPHRPNYFFIGNIALSLIAGVATNLLSGAT 62  
QY 100 TLKLSPALWFAREGGVFVALTASVLSLALALERSLTMAR---RGPAPVSSRGRTLAMAA 156  
DB 63 TYKLPAPQWFLREGSMFVALSVLSVSLALAIERYITMLKHLNG---SNNRFLLLIS 118  
QY 157 AAWGVSLILGLLPALGNCLGRDCACTVLPYAKAYVLCVL-APVGLAALCALYARI 215  
DB 119 ACWVSLILGLPGLWNCISALSSCTVPLPHKHILFCTLIIVFTLLLSIVILYCR 178  
QY 216 YQVQVANARLPARPAGTAGTSTRARRKPSLALRLTSLSVYLAFVACWGLFLLLLDV 275  
DB 179 YSLVTRSRRLTFKNI-----SKASRSSENVALTKVIVLSVFACWAPFLILLDV 233  
QY 276 APCARTCVLQADPEFLGAMANSLLNPIIYTLNDRHALRLVCCGRHSCGRDPGSG 335  
DB 234 GCKVTCIDLPRAEYELFVIAVINGTNPPIIYTLNEMRAFRIMCC-----KCPSG- 286  
QY 336 QQSAAEAGGLRCLPPLDGSFSGSRS--PQRDGLDTSGSTGSPG 383  
DB 287 -----DSACKFKRPPIAGMEFSRSKSDNSHPQDEGNPETINSSG 328

RESULT 10  
US-09-082-088-2  
; Sequence 2, Application US/09082088  
; Patent No. 6130067  
; GENERAL INFORMATION:  
; APPLICANT: TSUI, PING  
; TITLE OF INVENTION: HUMAN EDG3sb GENE  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ratner & Prestia  
; STREET: P.O. Box 980  
; CITY: Valley Forge  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19482  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082,088  
; FILING DATE: 20-MAY-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Prestia, Paul F  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GP-70453  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-407-0700  
; TELEFAX: 610-407-0701  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 378 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-082-088-2

Query Match 38.4%; Score 774.5; DB 4; Length 378;  
Best Local Similarity 44.7%; Pred. No. 2.9e-52;  
Matches 156; Conservative 55; Mismatches 109; Indels 29; Gaps 4;  
QY 6 LRPAPEVIVLHYNTGKRGARYOPGAGLRADAVVCAVLENLAVLLVLRHP 65  
DB 9 LQPVRGNETLREHYQVGLAGRLKEASEGSLTTTVLFLVICSFIVLENLAVLIAIKNN 68  
QY 66 RFPAPFLILGSLTSLDLAGAAYAAANILLGSLTSLKSPALWFAREGGVFVALTASVLS 125  
DB 69 KFNRMVFFIGNLALCDLAGIAYKVNILMSKKFSLPTVWFLREGSMFVALGASTCS 128  
QY 126 LLAIALERSUTMARRGPAPVSSRGRTLAMAAAANGVSLLLGLLPALGNCLGRDCACTV 185  
DB 129 LLAIALERHLTKMRPYDANKRHVFLILGMCWLIATLGAIPILGNCLHNPDCSTI 188  
QY 186 LPYAKAYVLCVLAPVGLAALCALYARIYCOVRANARLPARPCTAGTSTRARRKPR 245  
DB 189 LPYAKAYVLCVLAPVGLAALCALYARIYCOVRANARLPARPCTAGTSTRARRKPR 245  
QY 246 SLALLRTLVSVLAFVACWGLFLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPII 305  
DB 238 SMALLRTVIVSVFVACWGLFLILIDVACVQACPLFKAQWFLVILVLSAMNPVI 297  
QY 306 YTLTNDRLHALRLVCCGRHSC---GR-----DPSGSOOSAS 340  
DB 298 YTLASKEMRRAPFLVLC---NCLVGRGARASPIQALDPSRSKSSSS 342

RESULT 11  
US-08-196-989B-2  
; Sequence 2, Application US/08196989B  
; Patent No. 5585476  
; GENERAL INFORMATION:  
; APPLICANT: MacLennan, A. John  
; TITLE OF INVENTION: Molecular Cloning and Expression of  
; TITLE OF INVENTION: G-Protein Coupled Receptors  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Saliwanchik & Saliwanchik  
; STREET: 2421 N.W. 41st Street, Suite A-1  
; CITY: Gainesville  
; STATE: FL  
; COUNTRY: US  
; ZIP: 32606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/196,989B  
; FILING DATE: 15-FEB-1994

CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Lloyd, Jeff  
REGISTRATION NUMBER: 35,589  
REFERENCE/DOCKET NUMBER: MAC-100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
TELEFAX: 904-372-5800  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 352 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-196-989B-2

Query Match 36.3%; Score 732.5; DB 1; Length 352;  
Best Local Similarity 45.5%; Pred. No. 4.4e-49;  
Matches 166; Conservative 48; Mismatches 122; Indels 29; Gaps 6;

QY 3 SGLLRPAPSEVIVLHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIIVLENLAVLLVIG 62  
Db 6 SEYLNPEKVE---HNTYTKETLDMQETPSRKVASAFIIL--CCAIVVENLLVLIAYA 59  
QY 63 RHRPRHAPMFLLSGSLTSLDLAGAAYANILLSGPLTKLSPALWFAREGGVFVALTAS 122  
Db 60 RNSKFHSAMYLEFLGNLAASDLGAVAFVANTLLSGPVTLSLPLQWFARESAFITLSAS 119  
QY 123 VLSLLAIALERSITMARRGPAPVSSRGRTLMAAAAGVSLLLGLLPALGNWCLGRDAC 182  
Db 120 VFSLLAIAIEROVAIAKVKLYGSDKSCRMMLIGASWLSILGLPLGNWCLDHEAC 179  
QY 183 STVPLPYAKYVLCVLAFFVGLAALCALYARIYCOVRANARRLPARPGTAGTSTRARR 242  
Db 180 STVPLPYAKHYVLCVTFISVILLALVALYVRIYFVRS-----SHADVAG----- 225  
QY 243 KPSLALLRTLSVLLAFVACWGPFLFLLLDVACPARTCPVLLQADPFLGLAMANSLLN 302  
Db 226 -POTLALLKTVITVLGVFIICWLPAPAFSILLDSTCPVRACPVLYKAHYFFAFATLSLN 284  
QY 303 PIITYITNDRDLRHALLRLVCC---GRHSCGRDPSSQSSASAAEASGGLRCL-----PP 354  
Db 285 PVIYTWRSRDLRREVLRPLLCWRQKGATGRGNGPNHRLPLRSSSSSLRGLHMTSPT 344  
QY 355 GLDGS 359  
Db 345 FLEGN 349

RESULT 12  
US-08-760-936-2  
Sequence 2, Application US/08760936  
Patent No. 5856443  
GENERAL INFORMATION:  
APPLICANT: MacLennan, A. John  
TITLE OF INVENTION: Molecular Cloning and Expression of  
TITLE OF INVENTION: G-Protein Coupled Receptors  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik  
STREET: 2421 N.W. 41st Street, Suite A-1  
CITY: Gainesville  
STATE: FL  
COUNTRY: US  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/760,936  
FILING DATE: December 6, 1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Pace, Doran R.  
REGISTRATION NUMBER: 38,261  
REFERENCE/DOCKET NUMBER: MAC-100C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 352-375-8100  
TELEFAX: 352-372-5800  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 352 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-760-936-2

Query Match 36.3%; Score 732.5; DB 2; Length 352;  
Best Local Similarity 45.5%; Pred. No. 4.4e-49;  
Matches 166; Conservative 48; Mismatches 122; Indels 29; Gaps 6;

QY 3 SGLLRPAPSEVIVLHNTYTKLGRARYOPGAGLRADAVVCLAVCAFIIVLENLAVLLVIG 62  
Db 6 SEYLNPEKVE---HNTYTKETLDMQETPSRKVASAFIIL--CCAIVVENLLVLIAYA 59  
QY 63 RHRPRHAPMFLLSGSLTSLDLAGAAYANILLSGPLTKLSPALWFAREGGVFVALTAS 122  
Db 60 RNSKFHSAMYLEFLGNLAASDLGAVAFVANTLLSGPVTLSLPLQWFARESAFITLSAS 119  
QY 123 VLSLLAIALERSITMARRGPAPVSSRGRTLMAAAAGVSLLLGLLPALGNWCLGRDAC 182  
Db 120 VFSLLAIAIEROVAIAKVKLYGSDKSCRMMLIGASWLSILGLPLGNWCLDHEAC 179  
QY 183 STVPLPYAKYVLCVLAFFVGLAALCALYARIYCOVRANARRLPARPGTAGTSTRARR 242  
Db 180 STVPLPYAKHYVLCVTFISVILLALVALYVRIYFVRS-----SHADVAG----- 225  
QY 243 KPSLALLRTLSVLLAFVACWGPFLFLLLDVACPARTCPVLLQADPFLGLAMANSLLN 302  
Db 226 -POTLALLKTVITVLGVFIICWLPAPAFSILLDSTCPVRACPVLYKAHYFFAFATLSLN 284  
QY 303 PIITYITNDRDLRHALLRLVCC---GRHSCGRDPSSQSSASAAEASGGLRCL-----PP 354  
Db 285 PVIYTWRSRDLRREVLRPLLCWRQKGATGRGNGPNHRLPLRSSSSSLRGLHMTSPT 344  
QY 355 GLDGS 359  
Db 345 FLEGN 349

RESULT 13  
US-08-852-824-4  
Sequence 4, Application US/08852824C  
Patent No. 6060272  
GENERAL INFORMATION:  
APPLICANT: Li, et al.  
TITLE OF INVENTION: Human G-Protein Coupled Receptors  
FILE REFERENCE: 1488.122000  
CURRENT APPLICATION NUMBER: US/08/852,824C  
CURRENT FILING DATE: 1997-05-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 4  
LENGTH: 384  
TYPE: PRT  
ORGANISM: genomic  
US-08-852-824-4

Query Match 31.3%; Score 631.5; DB 3; Length 384;

```
; LIBRARY: TYNOR01
; CLONE:. 14690
; US-08-845-566-1
```

Query Match 27.1%; Score 548; DB 2; Length 509;  
Best Local Similarity 46.2%; Pred.No.9.6e-35;  
Matches 121; Conservative 37; Mismatches 84; Indels 20; Gaps 5;

QY	12 SEVIVLHYNTCKLRGARYPGCAGRADAVCYLCVAFIVLENLAVALLVGRHPRFAPHM 71   :         :   :       :   :           :   :   :
Dd	22 SELIXLHYNHSRGLAGRGGPEDGGALURGSUVAASCLVLVDENLLVLAATSHMSRRWV 81   :         :   :       :   :           :   :   :
QY	72 FILLGSITLSDLLGAAYAANILISGPLTKLSPALMFAREGGVFVALTASVLSLIAIAL 131 :   :                       :   :
Dd	82 YYCVLVNITSDDLTCNAVLANVLLSGATFFRLAPAQWFLEGLFTALAASTSLFTAG 141   :                       :   :
QY	132 EESLTWARGPAPVSRSRG-----RTLMAAAAANGVSLLLGLLPALGNCLGRDCASTVL 186   :
Dd	142 ERFATMV---PVAESGATKTSRYVGFIGLCWLLAALLGMPLLGWNCLCAFDCSSLL 197   :
QY	187 PLYAKAYVLFVCVIAFWGIIAAICALLYARIYCQVRANARRLPARPGTAGTTSTRARKPRS 246     :       :     :     :   :   :   :   :   :
Dd	198 PLYSKRYILFCVIFAGVLATTIMGLYAIFRLVQAOSQKAP-RPA-----ARRKAR-- 247     :       :     :     :   :   :   :   :   :
QY	247 LALLRTLSVVLLFAVACWGPLE 268     :   :       :
Dd	248 -RLKTVMILLFLAVCWGPLE 268     :   :       :

```

US-08-763-938-2
; Sequence 2, Application US/08763938
; Patent No. 6140060
; GENERAL INFORMATION:
; APPLICANT: CHUN, Jerold J.M.
; APPLICANT: HECHT, Jonathan H.
; TITLE OF INVENTION: CLONED LY SOPHOSPHATIDIC ACID
; TITLE OF INVENTION: RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray and Oram LLP
; STREET: 655 15th Street, N.W., Suite 330 - G St. Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/763,938
; FILING DATE: 12-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: JAHNS, Kristina M.
; REGISTRATION NUMBER: 41,092
; REFERENCE/DOCKET NUMBER: P8074-6018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-5000
; TELEFAX: (202) 638-4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 364 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-763-938-2

```

```

Query Match      25.7%; Score 518.5; DB 4; Length 364;
Best Local Similarity 34.1%; Pred. No. 1.2e-32;
Matches 115; Conservative 70; Mismatches 129; Indels 23; Gaps 5;

Qy 12 SEVIVLHNYTGKRGARYQP-----GAGLRADAVVCLAVCAFIIVLENLAVLLVLRHP 65
Db 27 NESIAFTYNRSGKYLATEWNTSKLVNGLG-----ITVCVFIMLANLLVNVAIYVNR 78

Qy 66 RFHAPMELLGSLTSLDLAAGAAANILSGPLTKLSPALAFAREGGVFVALTASVLS 125
Db 79 RFHEPIIYLANLAADFFAGLAYFYLMFTGPNTRRLIVSTWLLRQGLDTSLTASVAN 138

Qy 126 LLATALERSLTMARRGAPVSSRGRRTLMAAAAGVSLLLGLLPALGNCLGRDLACSTV 185
Db 139 LLATAIERHITVFRMQLHTRMSNRVVVIVITWTMAIVNGAIPSVGWNCICDIDHCSNM 198

Qy 186 LPLVAKAYVLCVLAFCVILAAICALYARYICOVRANARLPAAPCGTAGTTSTRARRKPR 245
Db 199 APLISDSYLVFWAIFNVTVVWVVLVFAHIFGYVORTWRMSRH-----SSGPRNRDT 252

Qy 246 SLALLRTLVSVLLAFVACWGFLELLLLLDVACPARTCPVLLQADPELGLAMANSILNPIT 305
Db 253 MMSLLKTWIVLVGAFIVCWTPGLVLLLDVCCP--QCDVLAYEKFFLLAEFNSAMNPIT 310

Qy 306 YLTNRDLRHALLRLVCCGRHSCGRDPS-GSQQSASA 341
Db 311 YSYRDKEMSATFRQILCCQRNENPNPGTEGSDRSASS 347

```

Search completed: January 16, 2002, 23:40:02  
Job time: 2159 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2002, 23:35:18 ; Search time 65.2 seconds

(without alignments)  
464.992 Million cell updates/sec

Title: US-09-842-316-2

Perfect score: 2019

Sequence: 1 MESGLLRPAVPSEVIVLHYN.....TGSPGAPTAARTIVSEPAD 398

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_68:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	853.5	42.3	383	2 I33870	Edg-1 orphan recep
2	834	41.3	381	2 A35300	G protein-coupled
3	815.5	40.4	362	3 JC7559	sphingosine 1-phos
4	774.5	38.4	378	2 JC5245	G protein-coupled
5	732.5	36.3	352	2 JC1465	probable G protein
6	506.5	25.1	364	2 JC5293	lysophosphatidic a
7	440.5	21.8	180	2 E48909	G protein-coupled
8	357	17.7	330	2 A55689	G protein-coupled
9	347	17.2	330	2 S40434	G protein-coupled
10	325.5	16.1	473	2 A33117	cannabinoid recep
11	318	15.8	362	2 I65990	G protein-coupled
12	310	15.4	472	2 S17595	cannabinoid recep
13	308	15.3	363	2 S48697	probable G protein
14	289.5	14.3	360	2 S36750	cannabinoid recep
15	282.5	14.0	372	2 I49008	melanocortin-5 rec
16	282	14.0	466	2 S36794	beta-1-adrenergic
17	281.5	13.9	325	2 I46416	melanocyte-stimula
18	278	13.8	325	2 JC5592	melanocortin 5 rec
19	275	13.6	325	2 JC2193	melanocortin recep
20	273.5	13.5	347	2 S70364	cannabinoid recep
21	273	13.5	332	2 A57055	melanocortin recep
22	272.5	13.5	464	2 S12591	beta-1-adrenergic
23	272	13.5	325	2 JN0764	melanocortin recep
24	267	13.2	480	2 I53053	beta 1 adrenergic
25	264	13.1	360	2 B46647	melanocortin recep
26	262	13.0	323	2 S36636	melanocortin recep
27	259.5	12.9	477	1 QRHUB1	beta-1-adrenergic
28	259	12.8	323	2 S43850	melanocortin 3 rec
29	259	12.8	428	2 A55044	beta-4C-adrenergic

30	257.5	12.8	440	2 JC5520	serotonin receptor
31	255.5	12.7	405	2 S65459	beta-3-adrenergic
32	252.5	12.5	374	2 I77467	serotonin receptor
33	250.5	12.4	389	2 S68422	serotonin receptor
34	247.5	12.3	320	2 S18444	G protein-coupled
35	247.5	12.3	386	2 S18637	serotonin receptor
36	246.5	12.2	386	2 S54153	serotonin receptor
37	246	12.2	317	2 T12055	melanocyte stimula
38	244.5	12.1	386	2 A42688	serotonin receptor
39	244.5	12.1	390	2 JN0268	serotonin receptor
40	243.5	12.1	418	2 G02953	beta-3-adrenergic
41	243.5	12.1	560	2 A38731	alpha-1A adrenergic
42	243	12.0	390	2 S58126	serotonin receptor
43	243	12.0	437	2 I57942	5-hydroxytryptamin
44	243	12.0	445	2 T42203	serotonin receptor
45	240	11.9	517	2 A45121	alpha-1B adrenergic

## ALIGNMENTS

RESULT 1

I53870

Edg-1 orphan receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 24-Sep-1999

C:Accession: I53870

R:Lado, D.C.; Browe, C.S.; Gaskin, A.A.; Borden, J.M.; MacLennan, A.J.

Gene 149, 331-336, 1994

A:Title: Cloning of the rat edg-1 immediate-early gene: expression pattern suggests d

A:Reference number: I53870; MUID:95047498

A:Accession: I53870

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-383 <RES>

A:Cross-references: EMBL:U0303; NID:g595396; PIDN:AAA83418.1; PID:g595397

C:Superfamily: G protein-coupled receptor edg-1

Query Match 42.3%; Score 853.5; DB 2; Length 383;  
Best Local Similarity 47.6%; Pred. No. 1.2e-59;  
Matches 177; Conservative 66; Mismatches 110; Indels 19; Gaps 5;

QY 13 EVIVLHYNTGKLRGARYQPGAGLRADAVCLVCAFTVLENLAVLVLRHPRFHPMF 72

DB 24 DIIVRHYNTGKLR-NIGVEKDHGKLTSEVFLICCLILENIFVLLTIWTKKFRPMY 82

QY 73 LLGSLTLDLLAGAAAYANILLGSLTKLSPALWFAREGGVFVALTSVLSLALALE 132

DB 83 YFIGNALSDLLAGVAYTANLLSGATTYKLTPAQWFLREGSMFVALSVFSLALAIIE 142

QY 133 RSLTWARRGPVPSRGRTLAWAAAANGVSLLLGLLPALGNCLGRDLACSTVLPYAKA 192

DB 143 RYITMLKMKLHNGSSRSRSLISACWVLSLILGLPTMGWNCISSLSGCTVLPYHKH 202

QY 193 VYLFVCLAVFGVTLAICALYARIYCOVRANARRLPARGTAGTSTPARKPRSLALLRT 252

DB 203 YILFTTFTLLLSLVILYICRIYSLVTRSRLLTFRK----NISKASRSSEKSLALLKT 258

QY 253 LSVLLAFVACWGPFLFLLLDLVACPARTCPVLQADPFPLGLAMANSLLNPIIYTTNRD 312

DB 259 VIIVLSVFIACWAPFLFLLLDVGCKARTCDILYKAEVFLVAVLNSGTNPYIITLNKE 318

QY 313 LRHALLRLVCCGRISGDRPSGQSQSASAPASGGLRCLPPLGLDGSFSGSERSS--PQD 371

DB 319 MRAPRII-----SCCKCPNG-----DSAGKFRPIIPGMEFSRSKSDNSHPQKD 365

QY 372 GLDTSGTSGSPG 383

DB 366 DGDNPETIMSSG 377

RESULT 2

A35300  
G protein-coupled receptor edg-1 - human  
C:Species: Homo sapiens (man)  
C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 24-Sep-1999  
C:Accession: A35300  
R:HLA, T., Maciag, T.  
J. Biol. Chem. 265, 9308-9313, 1990  
A:Title: An abundant transcript induced in differentiating human endothelial cells encoded by a human gene  
A:Reference number: A35300; MUID:90264425  
A:Accession: A35300  
A:Molecule type: mRNA  
A:Residues: 1-381 <HLA>  
A:Cross-references: GB:M31210; NID:g181948; PIDN:AAA52336.1; PID:g181949  
C:Genetics: GDB:ECGF1  
A:Cross-references: GDB:127754; OMIM:131222  
A:Map position: 22q13-22q13  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein

Query Match 41.3%; Score 834; DB 2; Length 381;  
Best Local Similarity 46.8%; Pred. No. 4e-57;  
Matches 174; Conservative 68; Mismatches 110; Indels 20; Gaps 5;

QY 13 EVIVLYNTGKLRGARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHAPMF 72  
DB 23 DIIVRYNTGKLNISADKENS-IKLTGVFLICCTFIENLAVLLVLRHPRHAPMF 81  
QY 73 LLLGSLTSLDLAAGAAANILLSPALPFAFGVGVFVLTALSLALE 132  
DB 82 YFIGNLALSDLLAGVAYTANLLSGATYKLTPTQWFLREGSMFVLSALSLALE 141  
QY 133 RSLTMARRGPAPVSSRGRTLAMAAANGVSLLLGLPALGWNCLGRDACSSTVLPYAKA 192  
DB 142 RYITMLKMLHNGSNFRLFLISACWVSLTGLGPIWGNICISALSGSTVLPYHKK 201  
QY 193 YVLCFLAFVGLAALCALYARIYCOVRANARLPARPCTAGTTSTRARRKPSLALLRT 252  
DB 202 YILFCVTFTLLSLVILYCHYSLVTRSRRLTFKNI-----SKARSENVALKLT 256  
QY 253 LSVLLAFVACGPIFLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPIYTTNRD 312  
DB 257 VIIVLSVFTACWAPLFIILLDVGKVKTCDFLFAEYFLVAVLNSGTPNIYITLTK 316  
QY 313 LRHALLRLVCCGRHSGRDPGSGQSASAAEAGGLRCLPPGLDGSFSGSERSS-PQRD 371  
DB 317 MRRAPTRIM-----SCCKCPG-----DSAGKFRPIIAGMEFSRKSNDSSHPQKD 363  
QY 372 GLDTSGSTGSPG 383  
DB 364 EGDNPETIMSSG 375

RESULT 3  
JC7559  
sphingosine 1-phosphate receptor - zebra fish  
N:Alternate names: endothelial differentiation gene 1 receptor  
C:Species: Brachydanio rerio (zebra fish)  
C:Date: 31-Mar-2001 #sequence\_revision 31-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: JC7559  
R:Im, D.S.; Ungar, A.R.; Lynch, K.R.  
Biochem. Biophys. Res. Commun. 279, 139-143, 2000  
A:Title: Characterization of a zebrafish (Danio rerio) sphingosine 1-phosphate receptor  
A:Reference number: JC7559; MUID: 20563813  
A:Accession: JC7559  
A:Molecule type: DNA  
A:Residues: 1-362 <IMA>  
A:Cross-references: GB:A3321294  
C:Comment: This receptor, a G protein-coupled receptor, mediates sphingosine 1-phosphate or also functions in the developing central nervous system in cell proliferation, apoptosis  
C:Genetics:  
A:Gene: edg1

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 40.4%; Score 815.5; DB 3; Length 362;  
Best Local Similarity 47.1%; Pred. No. 1e-55;  
Matches 172; Conservative 59; Mismatches 107; Indels 27; Gaps 7;

QY 11 VSEIVLHYNTGKLRGARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPRHAP 70  
DB 1 MODLLARHYNFTGFRKVKHDP--GLKADSVVFIIVCCFIENLAVLLVLRHPRHAP 58  
QY 71 MFLLLGSLTSLDLAAGAAANILLSPALPFAFGVGVFVLTALSLALE 130  
DB 59 MYFIGNLALSDLLAGVYVYTNILLSGANTYKLTPTQWFFREGSMFVALAASVFLSLAIA 118  
QY 131 LERSLTMARRGAPVSSRGRT--LAWAAAGVSLLLGLPALGWNCLGRDACSSTVLP 187  
DB 119 IERHUTMK---MKLHNGKTCRVFMLISTVWFIAIILGLLPVWGNICIDSNINCSTVLP 175  
QY 188 LYAKAYVLCFLAFVGLAALCALYARIYCOVRANARLPARPCTAGTTSTRARRKPSL 247  
DB 176 LYHKAYILFCITVFSVILMAIVLYARIYALVTRSRKLVPRKVGANGSKSEK--SM 233  
QY 248 ALLRTLSVLLAFVACGPIFLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPIYT 307  
DB 234 ALLKTVIIVLSCTACWAPLFIILLDVGKVKTCDFLFAEYFLVAVLNSAMNPLIYT 293  
QY 308 LTRDLRLVCCGRHSGRDPGSGQSASAAEAGGLRCLPPGLDGSFSGSERSS 367  
DB 294 LLSNEMRAPIKMLNCG-----VCVQPSGKFSRPI-MGAEFSTKSDNS 337  
QY 368 -PQRD 371  
DB 338 HENKD 342

RESULT 4

JC5245  
G protein-coupled receptor - human  
C:Species: Homo sapiens (man)  
C:Date: 13-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: JC5245  
R:Yamaguchi, F.; Tokuda, M.; Hatase, O.; Brenner, S.  
Biochem. Biophys. Res. Commun. 227, 608-614, 1996  
A:Title: Molecular cloning of the novel human G protein-coupled receptor (GPCR) gene  
A:Reference number: JC5245; MUID:97032811  
A:Accession: JC5245  
A:Molecule type: DNA  
A:Residues: 1-378 <YAM>  
A:Cross-references: EMBL:X83864; NID:g1770395; PIDN:CAA56744.1; PID:g1770396  
C:Comment: This protein transduces hormone and neurotransmitter signals into intracellular  
C:Genetics:  
A:Gene: edg-3  
A:Map position: 9q22.1-q22.2  
C:Superfamily: G protein-coupled receptor edg-1  
C:Keywords: G protein-coupled receptor

Query Match 38.4%; Score 774.5; DB 2; Length 378;  
Best Local Similarity 44.7%; Pred. No. 1.5e-52;  
Matches 156; Conservative 55; Mismatches 109; Indels 29; Gaps 4;

QY 6 LRPAPVSEIVLYHYNTGKLRGARYOPGAGLRADAVVCLAVCAFIENLAVLLVLRHPR 65  
DB 9 LQVRGNETLREHYQYVGLKAGRLKEASEGSTLTFTVFLVICSFIENLAVLLVLRHPR 68  
QY 66 RFHAPMFLLLGSLTSLDLAAGAAANILLSPALPFAFGVGVFVLTALSLALE 125  
DB 69 KFNRMVFFIGNLALCDLLAGIAYKVNLGKTKFSLSPVWFLREGSMFVALGASTCS 128  
QY 126 LLATALERSLTMARRGAPVSSRGRTLAMAAANGVSLLLGLPALGWNCLGRDACSSTV 185  
DB 129 LLATAIERHUTMKRPRYDANKRHRVFLIIGMCMWLTFTAGLPILGWNCLHNLPCSTY 188



Db 345 FLEGN 349

RESULT 6

JC5293

lysophosphatidic acid receptor - human

N:Alternate names: Edg2 protein

C:Species: Homo sapiens (man)

C>Date: 16-Apr-1997 #sequence\_revision 09-May-1997 #text\_change 10-Dec-1999

C:Accession: JC5293

R:An, S.; Dickens, M.A.; Bleu, T.; Hallmark, O.G.; Goetzl, E.J.

Biochem. Biophys. Res. Commun. 231, 619-622, 1997

A:Title: Molecular cloning of the human Edg2 protein and its identification as a func

A:Reference number: JC5293; MUID:97224397

A:Contents: lung

A:Accession: JC5293

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-354 <ANA>

C:Cross-references: GB:U08011; NID:g1857424; PIDN:AAC51139.1; PID:g1857425

C:Superfamily: G protein-coupled receptor edg-1

Query Match 25.1%; Score 506.5; DB 2; Length 364;

Best Local Similarity 32.9%; Pred. No. 7.3e-32;

Matches 112; Conservative 73; Mismatches 130; Indels 25; Gaps 5;

QY 12 SEVILVHYNTGKLRGARYOP-----GAGLRADAVCVLAVCAFTVLENLAVLVLRGHP 65

Db 27 NESIAFYNRSGKHLATENTVSKLVMLG-----ITVCIFIMLANLLVMVAIVNR 78

QY 66 RHAPFLLGSLTSDLLAGAAAYANILLSPGLTKLSPALWFAREGGVFALTASVLS 125

Db 79 REHFPYILMANLAADFFAGLAYFYLMENTGPNTRRLTVSTWLLRGLDITSLTASVAN 138

QY 126 LLAIALERSLTVARRGPVSSRGRTLMAAAAGVSLGLLGLPALGNCLGRDLDCSTV 185

Db 139 LLAIALERITVFMQLTRMSNRVVVVVITMTAIVMGAIPSVGMNCICDIENCSNM 198

QY 186 LPLYAKAYVLCVLAIFVGLAALCAIYARYICQVRANARRLPARGTAGTTSVRRKPR 245

Db 199 APLYSDSYLVFWAIFENLVTFVVMVLYAHIFGYVQRTRMSRH-----SSGPRNRDT 252

QY 246 SLALLRTLSVLLAFVACWGPFLFLLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPII 305

Db 253 MNSLFTVVIVLGAFTIICWTGELVLLLDVCCP--QCDVLAYERFFLLAEFSANMPII 310

QY 306 YTLTNRDLRHALLRLVCCGRHSCGRDPSGSOOSASAAEAS 345

Db 311 YSVROKENSATRFQILCCQR---SENPTGPTESDRSASS 347

RESULT 7

E48909

G protein-coupled receptor Gpcrl3 - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 01-Dec-1995 #sequence\_revision 01-Dec-1995 #text\_change 19-May-2000

C:Accession: E48909

R:Wilkie, T.M.; Chen, Y.; Gilbert, D.J.; Moore, K.J.; Yu, L.; Simon, M.I.; Copeland,

Genomics 18, 175-184, 1993

A:Title: Identification, chromosomal location, and genome organization of mammalian G

A:Reference number: A48909; MUID:94116980

A:Accession: E48909

A:Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-180 <WIL>

C:Cross-references: GB:L20334; NID:g438786; PIDN:AAA16846.1; PID:g438787

C:Superfamily: G protein-coupled receptor edg-1

C:Keywords: G protein-coupled receptor

Query Match 21.8%; Score 440.5; DB 2; Length 180;

Best Local Similarity 50.5%; Pred. No. 4.6e-27;

QY 186 LPLYAKAYVLCVLAIFVGLAALCAIYARYICQVRANARRLPARGTAGTTSVRRKPR 245

Db 189 LPLYSKYIAFCISIFTAILVTIVILYARIYFLVKSSRKV-----ANHNNNSR 237

QY 246 SLALLRTLSVLLAFVACWGPFLFLLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPII 305

Db 238 SMALLRTVIVSVFIACWSPFLFLFDVACRVQACPIFLFKAQWFIYVLAVLNSANMPTI 297

QY 306 YTLTNRDLRHALLRLVCCGRHSC---GR-----DPSGSOOSAS 340

Db 298 YTLASKEMRRAFFRLVC---NCLVRGRGARAPIQALDPSRSKSSSS 342

RESULT 5

JC1465

probable G protein-coupled receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 21-Jul-2000

C:Accession: JC1465

R:Okazaki, H.; Ishizaka, N.; Sakurai, T.; Kurokawa, K.; Goto, K.; Kumada, M.; Takuwa, Y.

Biochem. Biophys. Res. Commun. 190, 1104-1109, 1993

A:Title: Molecular cloning of a novel putative G protein-coupled receptor expressed in t

A:Reference number: JC1465; MUID:93176155

A:Accession: JC1465

A:Molecule type: mRNA

A:Residues: 1-352 <OKA>

A:Cross-references: GB:AB016931; NID:g3445557; PIDN:BAA32454.1; PID:g3445558

A:Experimental source: aortic smooth muscle

C:Superfamily: G protein-coupled receptor edg-1

C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot

F:35-59/Domain: transmembrane #status predicted <TM1>

F:67-95/Domain: transmembrane #status predicted <TM2>

F:110-128/Domain: transmembrane #status predicted <TM3>

F:148-173/Domain: transmembrane #status predicted <TM4>

F:190-210/Domain: transmembrane #status predicted <TM5>

F:234-255/Domain: transmembrane #status predicted <TM6>

F:272-293/Domain: transmembrane #status predicted <TM7>

F:19/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:142,145,218,219,329,330,331,332/Binding site: phosphate (Ser) (covalent) #status predi

F:313/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 36.3%; Score 732.5; DB 2; Length 352;

Best Local Similarity 45.5%; Pred. No. 2.5e-49;

Matches 166; Conservative 48; Mismatches 122; Indels 29; Gaps 6;

QY 3 SGLLRAPVSEVILVHYNTGKLRGARYOPGAGLRADAVCVLAVCAFTVLENLAVLVLRG 62

Db 6 SEYNLEPKVQE-----HYNTKETLDQETPPSRKVASAFIIL--CCATVWENLVLLIAVA 59

QY 63 RHPRFHPFLLGSLTSDLLAGAAAYANILLSPGLTKLSPALWFAREGGVFALTAS 122

Db 60 RNSKFSHSMYFLGNLAASDLAGAVFANTLLSGVLSLTPQWFAREGSAFITUSAS 119

QY 123 VLSLLAIALERSLTVARRGPVSSRGRTLMAAAAGVSLGLLGLPALGNCLGRDLDC 182

Db 120 VFSLLAIAIEROVAIAKRVLYGSDKSRMLIGASWLSLITLGLPLTGNWCLDHEAC 179

QY 183 STVPLYAKAYVLCVLAIFVGLAALCAIYARYICQVRANARRLPARGTAGTTSVRR 242

Db 180 STVPLYAKHYVLCVVTFISVLLAIALYIYRIYFVRS-----SHADVAS----- 225

QY 243 KPSRLALLRTLSVLLAFVACWGPFLFLLLLLDVACPARTCPVLLQADPFLGLAMANSLLN 302

Db 226 -POTLALLKTIVLGVFTICWLPFAFSLIDSTCPVACPVLYKHAHYFAFAFLNSELN 284

QY 303 PIYITLTNRDLRHALLRLVCC---GRHSCGRDPSGSOOSASAAEASGLRRCL-----PP 354

Db 285 PVIYTWRSNDLRREVLRPLLCWROQKGATGRGGNPGHRLPLLRSSSSSLERGLHMTPT 344

QY 355 GLDGS 359

:-:-:-

[illegible]

RESULT 8  
A53689  
G protein-coupled receptor 3 - human  
N:Alternate names: G protein-coupled receptor GPR3; orphan G-protein-coupled receptor AC  
C:Species: Homo sapiens (man)  
C>Date: 03-Mar-1995 #sequence\_revision 03-Mar-1995 #text\_change 19-May-2000  
C:Accession: A55689; S58521; C55733  
R:Alisma, T.P.; Klefer, J.; Liu, M.L.; Baker, E.; Sutherland, G.R.; Shine, J.  
Genomics 24, 391-394, 1994  
A:Title: Isolation and chromosomal localization of a novel human G-protein-coupled receptor  
A:Reference number: A55689; MUID:95213036  
A:Accession: A55689  
A:Status: Preliminary  
A:Molecule type: DNA; mRNA  
A:Residues: 1-330 <IIS>  
R:Cross-references: GB:L32830; GB:L32831; NID:g602311; PID:AAAT35560.1; PID:g602312  
R:Eggarikokx, D.; Denet, J.F.; Labbe, O.; Hayashi, Y.; Refetoff, S.; Vassart, G.; Parmant  
Biochem. J. 309, 837-843; 1995  
A:Title: Molecular cloning of an orphan G-protein-coupled receptor that constitutively a  
A:Reference number: S58521; MUID:95366960  
A:Accession: S58521  
A:Status: Preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 1-330 <EGG>  
R:Cross-references: GB:X83956; NID:g1061125; PIDN:CAA58787.1; PID:g1061126  
R:Marchese, A.; Docherty, J.M.; Nguyen, T.; Heiber, M.; Cheng, R.; Heng, H.H.Q.; Tsui, I.  
Genomics 23, 609-618, 1994  
A:Title: Cloning of human genes encoding novel G protein-coupled receptors.  
A:Reference number: A55733; MUID:95154831  
A:Accession: C55733  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-292 <MAR>  
R:Cross-references: GB:U13668; NID:g577416; PIDN:AAA64594.1; PID:g577417  
C:Genetics:  
A:Gene: GDB:GPR3  
A:Cross-references: GDB:371695; OMIM:600241  
A:Map position: lp36.1-ip34.3  
C:Superfamily: melanocortin receptor  
C:Keywords: G protein-coupled receptor; transmembrane protein

```

Query Match      17.7%; Score 357; DB 2; Length 330;
Best Local Similarity 35.5%; Pred. No. 2.3e-20;
Matches 107; Conservative 40; Mismatches 108; Indels 46; Gaps 12;

QY 39 DAVGLVCAFTVLENTAVLVVGRHPRFHAFMFLGLSGLTSLDLAAG---AAYAAINTLL 95
| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 44 DVVLICIS-GLTVSCENALVVAIIVTGTFAPRAPMFLVLVGLSVLADVLGLGLVHLHFAAVFC 102
| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|

QY 96 SGPLTKLSPALWTFAREGGVFVALTASVLSLALAIERSLUTWARR-----GPAPVSSRRGT 151
| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB 103 IGSAGMSL-----VLVGLVLAFTAGISGLLAIIVDRLVSLNALTYYSETVTV---RT 153

```

```

QY 152 LAMAAAAGVSLLLGLPALGWNCLGBLDACSTVLPYAKAYVLCVLAFAVGLAAICAL 211
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 154 YVMLALVWGGAIGLGLLPVLAWNCLDGLTTCGVVYPL-SKNHLVVLAIAPFMVFGIMQL 212
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 212 YA---RIYC-----QVRANARLPPARPGTAGTTSTRARRKPRSLLRLTSLVVLAFVACW 264
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 YAQICRIVCRHAOQIALQRHLLPA-----SHVAVTRKG-----IATLAVVLGAFACW 260
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 265 GPFLFLLLLDVA--CPARTCFVLQADPFLGLAMANSLLNPITYTLNDRDLRHALLRLVC 322
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 261 LPFTVYCLLDGAHSPPLYTYTLTP-----ATYNSWNPITYAFRQDVQKVLWAVCC 313
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 323 C 323
      :
Db 314 C 314

RESULT 9
S40454
G protein-coupled receptor GPCR21 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 19-May-1994 #sequence_revision 10-Nov-1995 #text_change 20-Jun-2000
C/Accession: S40454
R/Saeki, Y.; Ueno, S.; Mizuno, R.; Nishimura, T.; Fujimura, H.; Nagai, Y.; Yanagihara
FEBS Lett. 336, 317-322, 1993
A/Title: Molecular cloning of a novel putative G protein-coupled receptor (GPCR21) wh
A/Reference number: S40454; MUID:94085630

```

A; Molecule type: mRNA  
A; Residues: 1-330 <SAE>  
A; Cross-references: GB:D21062; NID:g455487; PIDN:BA04641.1; PID:g455488  
C; Superfamily: melanocortin receptor  
C; Keywords: G protein-coupled receptor; phosphoinositide transmembrane protein

Query Match 17.2%; Score 347; DB 2; Length 330;  
Best Local Similarity 34.3%; Pred. NO. 1.4e-19;  
Matches 103; Conservative 43; Mismatches 110; Indels 44; Gaps 11;

[illegible]

RESULT 10  
A33117  
cannabinoid receptor CB1 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 05-Nov-1999  
C:Accession: A33117; C55879  
R:Watsuda, Y.A.; Lolait, S.J.; Brownstein, M.J.; Young, A.C.; Bonner, T.I.  
Nature 346, 561-564, 1990  
A:Title: Structure of a cannabinoid receptor and functional expression of the cloned  
A:Reference number: A33117; MUID:90332039  
A:Accession: A33117  
A:Molecule type: mRNA

A;Residues: 1-473 <MAT>  
A;Cross-references: GB:55812; NID:gl552375; PIDN:CAA39332.1; PID:g57249  
R;Shire, D.; Carillion, C.; Kaghad, M.; Calandra, B.; Rinaldi-Carmona, M.; Le Fur, G.; Car-  
J. Biol. Chem. 270, 3726-3731. 1995  
A;Title: An amino-terminal variant of the central cannabinoid receptor resulting from al-  
A;Reference number: A55879; MUID:95181329  
A;Accession: C55879  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: mRNA  
A;Residues: 1-107 <SHI>  
C;Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match: 16.1%; Score 325.5; DB 2; Length 473;  
Best Local Similarity 30.1%; Pred. No. 8.9e-18;  
Matches 91; Conservative 59; Mismatches 117; Indels 35; Gaps 9;

	QY	38	ADAVVCLVAFCIVILENAVLVLGRPHRPHA-----PMFLLLGSLSTLSDLLAGAAAYAA	91
	Dd	119	AIAVLSTLTGTFTVENLLVCVI-----LHRSRLCRPSYHFIGSLAVADDLGSVFVV	173
	QY	92	NILLSGPLTKLSPALWFAREGGVFALTASVLSILATALERSLTMARRGPAPVS-----	146
	Dd	174	SFDVDFHFKDSENVELFKLGVTASTTASVGSLTFATAIDRYISIHRR-----PLAYKRV	229
	QY	147	SRGRTLAWAAAAWGVSLLLGLPALGWNCNLGRDLACSTVLPYAKAVFLVCVLAFGILA	206
	Dd	230	TRPKAVAFCLMWTIAITVAILPGLGNCKKLQSVDIFPIDETYMFWIGVTSVLL	289
	QY	207	AICALYARIYCOVRANARLRPPAGTAG---TS-----TBARKPRSLALLTLVS	255
	Dd	290	FIVYAYWYLWKASHAVRIQR-GTOKSIIITHSEDKGVQVTPDOARMDIRLATLV	348
	QY	256	VLLAFVACHGPFLLLLLOVACPARTCPVLOADPFLG-LAMANSLNPIIYTLTNRDR	314
	Dd	349	IIVWLITCWGPLAIMIYYDYF--GKNMKLIKTVPFACSMCLLNSTVNPIIYALRSKDUR	406
	QY	315	HA	316
	Dd	407	HA	408

RESULT 11

I65990

G protein-coupled receptor 6 - human

C.Species: Homo sapiens (man)

C.Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 24-Sep-1999

C.Accession: I65990

R.Helber, M.; Docherty, J.M.; Shah, G.; Nguyen, T.; Cheng, R.; Heng, H.H.Q.; Marchese, A.

DNAA Cell Biol. 14, 25-35, 1995

A.Title: Isolation of three novel human genes encoding G protein-coupled receptors.

A.Reference number: I53033; MUID:95134353

A.Accession: I65990

A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA

A.Residues: 1-362 <RES>

A.Cross-references: GB:L36150; NID:g598156; PIDN:AAA63181.1; PID:g598157

C.Genetics:

A.Gene: GDB:GPR6

A.Cross-references: GDB:371713; OMIM:600553

A.Map position: 6q21-6q21

C.Superfamily: melanocortin receptor

Query Match	15.8%	Score 318;	DB 2;	Length 362;
Best Local Similarity	34.1%;	Pred. No. 2.6e-17;		
Matches 104;	Conservative 38;	Mismatches 125;	Indels 38;	Gaps 9;

```

32 PGAGLR-----DAVVCVAACAFTVENLAVLLVLGRPHFAPMFELLGSLTLDLAGA 87
||| || | : |:| || | | | | | | | | | | | |
65 PGLLPVNPMDVLLCVS-GTVIAGENALVVALLTAASPALTPMFLVGSLATADILACC 123
|||||
88 AYAANILLSPGLTKLSPALWFAREGGVFVALTSVLSLIAIALERSLTWMRRGPVPSS 147

```

```

Db      124 GLILHFVQ---YLPVSETVSLLTVGFLVASFAVSGLLAITVDRLSYLN---ALTTY 177
QY      148 RGRPL-----AMAAANGVSLLLGLLPALGWNCGLRLDACSTVLPYAKAVLFCVLAFIG 203
Db      178 SRTLLGVHLLAATVTSGLGLPVLGWNCIAERAACSVWRPL-ARSHVALLSAAFFM 236
QY      204 ILAAICALYARIYCOVRANRRLPARPCTAGTSTARRKPRSLALLRTLSVLLAFVAC 263
Db      237 VFGINLHLYRICOVWVRHAHQIALOQHCLAPPHLAATRKG-----VGTLLAVLGTFGAS 291
QY      264 WGPLFLLLLLDVACPARTCPVLLQADE-----FLGLAMANSLLNPIYTLTNRDLRHAL 317
Db      292 WLFF-----ALYCVGVSHEDPAVITYATLLPATYNSMINPIIYAFRNOETORAL 340
QY      318 LRLVC 322
Db      341 WLILC.345

RESULT 12
SI7595
cannabinoid receptor CB1 - human
C:Species: Homo sapiens (man)
C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C:Accession: SI7595; SI3668; A55879
R:Gerard, C.M.; Mollereau, C.; Vassart, G.; Parmentier, M.
Biochem. J. 279, 129-134, 1991
A:Title: Molecular cloning of a human cannabinoid receptor which is also expressed in
A:Reference number: SI7595; MUID:92028798

```

RESULT 12

S17595  
 cannabinoid receptor CBI - human  
 C:Species: Homo sapiens (man)  
 C:Date: 03-Feb-1994 #sequence\_revision 03-Feb-1994 #text\_change 21-Jul-2000  
 C:Accession: S17595; S13668; A55879  
 R:Gerard, C.M.; Mollereau, C.; Vassart, G.; Parmentier, M.  
 Biochem. J. 279, 128-134, 1991  
 A:Title: Molecular cloning of a human cannabinoid receptor which is also expressed in  
 A:Reference number: S17595; M01D:92028798  
 A:Accession: S17595  
 A:Molecule type: mRNA  
 A:Residues: 1-472 <GER1>  
 A:Cross-references: EMBL:X54937; NID:g29914; PIDN:CAA38699.1; PID:g29915  
 R:Gerard, C.; Mollereau, C.; Vassart, G.; Parmentier, M.  
 Nucleic Acids Res. 18, 7142, 1990  
 A:Title: Nucleotide sequence of a human cannabinoid receptor cDNA.  
 A:Reference number: S13668; M01D:91088303  
 A:Accession: S13668  
 A>Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-472 <GER2>  
 A:Cross-references: EMBL:X54937; NID:g29914; PIDN:CAA38699.1; PID:g29915  
 R:Shire, D.; Carillon, C.; Kaghad, M.; Calandra, B.; Rinaldi-Carmona, M.; Le Fur, G.;  
 J. Biol. Chem. 270, 3726-3731, 1995  
 A:Title: An amino-terminal variant of the central cannabinoid receptor resulting from  
 A:Reference number: A55879; M01D:95181329  
 A:Accession: A55879  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-124 <SH1>  
 A:Cross-references: GB:X81120  
 C:Genetics:  
 A:Gene: GDB:CNRI; CNR  
 A:Cross-references: GDB:127354; OMIM:114610  
 A:Map position: 6q14-q15  
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein  
 F:117-142/Domain: transmembrane #status predicted <TM1>  
 F:155-175/Domain: transmembrane #status predicted <TM2>  
 F:188-212/Domain: transmembrane #status predicted <TM3>  
 F:233-256/Domain: transmembrane #status predicted <TM4>  
 F:275-299/Domain: transmembrane #status predicted <TM5>  
 F:345-365/Domain: transmembrane #status predicted <TM6>  
 F:378-399/Domain: transmembrane #status predicted <TM7>  
 F:77,83/Binding site: carbohydrate (Asn) #status predicted

```
Query Match      15.4% Score 310; DB 2; Length 472;
Best Local Similarity 27.8%; Pred. No. 1.4e-16;
Matches 96; Conservative 65; Mismatches 142; Indels 42; Gaps 11;
```

Db 118 AIAVLSLTGTFVLENLVLVI-----LHRSLECRPSYHFGISLAVADLLGSVIFY 172  
 Qy 92 NILSGPLTKLSPALWFAREGGVVALFASVLSLALALERSLTMARRGPAPVS----- 146  
 Db 173 SFIDFHFHRKSRNVEFLKLGVTASFASVGSFLTAIDRYSIHR-----PLAYKRIV 228  
 Qy 147 SRGRTLMAAAANGSVSLGLLPLGAGNCLGRDACSTVLPYAKAYVLCVLAFAVGILA 206  
 Db 229 TRPKAVAFCLMTIIVIAVPLPLGWNCEKLSQVCSDFPHIDITYLFWIGVTSVLL 288  
 Qy 207 AICALYARIYCOVRANARLPARPAGT-----TS-----TRARRKPRSLALLRTLSV 255  
 Db 289 FIVAYMYLWKAHSHAVRMQR-GTQKSIITHTSEDGKVQVTRPDQARMDIRLAKTIVL 347  
 Qy 256 VLLAFVACGPLELLLLDVLVACPARTCPVLLQADPFLG-LAMANSLLNPIIYTTNDRDLR 314  
 Db 348 ILVVLICWGFLAIVDYF--GRMKLIKTVFAFCMLCLNLTNVPPIIYALRSKDLR 405  
 Qy 315 HALLRLV--CCGR-----HSCGRDPSSGQSSASAAEASGGLRRCL 352  
 Db 406 HAFRSMFPCGTAQPLDMSMGSDCLHKHANNAASVHRAAESCI 450

RESULT 13  
 S48697  
 probable G protein-coupled receptor protein - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 24-Sep-1999  
 C:Accession: S48697  
 R:Song, Z.H.; Young III, W.S.; Brownstein, M.J.; Bonner, T.I.  
 FEBS Lett. 351, 375-379, 1994  
 A:Title: Molecular cloning of a novel candidate G protein-coupled receptor from rat brain  
 A:Reference number: S48697; MUID:94364507  
 A:Accession: S48697  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-363 <SON>  
 A:Cross-references: EMBL:U012006; NID:9551333; PIDN:AAA21870.1; PID:9551334  
 C:Superfamily: melanocortin receptor  
 C:Keywords: G protein-coupled receptor

Query Match 15.3%; Score 308; DB 2; Length 363;  
 Best Local Similarity 32.9%; Pred. No. 1.5e-16;  
 Matches 100; Conservative 40; Mismatches 126; Indels 38; Gaps 9;

Qy 34 AGRADAV---VCLAVCAFIYL-ENLAVLLVGRHPRHAPMELLLGSLTSLDLAGAA 88  
 Db 66 SGLLSAVNPWDVLCVSGTVIAGENALVVALIATSTPALRTPMFLVGLSATADLLAGCG 125  
 Qy 89 YAANILLGPTLKLSPALWFAREGGVVALFASVLSLALALERSLTMARRGPAPVSSR 148  
 Db 126 LILHFVQYVVP---SETVSLIMVGLVAFASVSSLLAITVDYLSLYN---ALTYYS 179  
 Qy 149 GRTL-----AMAAANGSVSLGLLPLGAGNCLGRDACSTVLPYAKAYVLCVLAFAVGI 204  
 Db 180 RFTLLGVHLLATVTSVGLGLPLVGLNCLADRSQVVRPL-TKSHVALLSTFFVV 238  
 Qy 205 LAAICALYARIYCOVRANARLPARPAGTGTSTRARRKPRSLALLRTLSVLLAFVACW 264  
 Db 239 FGIMLHLYVRICQVVRHQAIOALQOCHLAPPLAATRKG-----VGTAVLVLTGTSASW 293  
 Qy 265 GPLFLLLLLDVACPARTCPVLLQADPFLG-----AMANSLLNPIIYTTNDRHALL 318  
 Db 294 LFP-----AIYCVGSGQEDPAIYTYATLLPATYNSMINPIIYAPRQEIQRALW 342  
 Qy 319 RLVC 322  
 Db 343 LLFC 346

RESULT 14  
 S36750

cannabinoid receptor CB2 - human  
 N:Alternate names: cannabinoid receptor, peripheral  
 C:Species: Homo sapiens (man)  
 C:Date: 22-Jan-1994 #sequence\_revision 10-Nov-1995 #text\_change 17-Mar-1999  
 C:Accession: S36750  
 R:Munro S.; Thomas, K.L.; Abu-Shaar, M.  
 Nature 365, 61-65, 1993  
 A:Title: Molecular characterization of a peripheral receptor for cannabinoids.  
 A:Reference number: S36750; MUID:93368659  
 A:Accession: S36750  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-360 <MUN>  
 A:Cross-references: EMBL:X74328  
 C:Superfamily: melanocortin receptor

Query Match 14.3%; Score 289.5; DB 2; Length 360;  
 Best Local Similarity 29.3%; Pred. No. 4.1e-15;  
 Matches 95; Conservative 50; Mismatches 150; Indels 29; Gaps 11;

Qy 38 ADAVVCLAVCAFIYLENLAVL-LVLGRHPRHAPMELLLGSLTSLDLAGAAANILL 96  
 Db 35 AVAVLCTLLGLLSALENVAVLYLSSHOLRRKPSYLFIGSLAGADELASFVACSFVNF 94  
 Qy 97 GPLTKLSPALWFAREGGVVALFASVLSLALALERSLTMARRGPA--PVSSRGRTLAM 154  
 Db 95 HVFHVDSKAVFLKIGSVTMTFTASVGSLLTATDRYLCL-RYPPSYKALLTRGRLVT 153  
 Qy 155 AAAANGSVSLGLLPLGAGNCLGRDACSTVLPYAKAYVLCVLAFAVGLAAICALYAR 214  
 Db 154 LGIMVLSALVSLPLMGWTCPPR--PCSELFPLIPNDYLLSNLLFTAFIFSGIITYGH 211  
 Qy 215 IYCOVRANARLPAR-----PGTAGTSTRARRKPRSLALLRTLSVLLAFVACWGPFL 269  
 Db 212 VLMKAHOHVASLSGHODRQVPGMA-----RMRLDVRLAKTGLGLVALLICWFPVLA 263  
 Qy 270 LILLDVACPARTCPVLLQADPFLG-LAMANSLLNPIIYTTNDRDLR---HALLRLVCCG 324  
 Db 264 LMAHSLA--TTLSDOVKKAFACFCSMLCLNLSMNVPVIYALRSSEIRSSAHCLAHWKCV 321  
 Qy 325 RHSCGRD--PSGQSSASAAEASG 346  
 Db 322 R-GLGSEAKEAPRSSVTETADG 344

RESULT 15  
 I49008  
 melanocortin-5 receptor - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 24-Sep-1999  
 C:Accession: I49008; A54245; J02244  
 R:Fathi, Z.; Iben, L.G.; Parker, E.M.  
 Neurochem. Res. 20, 107-113, 1995  
 A:Title: Cloning, expression, and tissue distribution of a fifth melanocortin receptor  
 A:Reference number: I49008; MUID:95258173  
 A:Accession: I49008  
 A:Status: preliminary; translated from GB/EMBL/DDBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-372 <RES>  
 A:Cross-references: EMBL:U08354; NID:9522165; PIDN:AAA76585.1; PID:9522166  
 R:Labbe, O.; Desarnaud, F.; Eggerickx, D.; Vassart, G.; Parmentier, M.  
 Biochemistry 33, 4543-4549, 1994  
 A:Title: Molecular cloning of a mouse melanocortin 5 receptor gene widely expressed 1  
 A:Reference number: A54245; MUID:94213827  
 A:Accession: A54245  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 48-120, F', 122-372 <LAB>  
 A:Cross-references: GB:X76295; NID:9498973; PIDN:CAA53943.1; PID:9498974  
 A:Experimental source: clone HGMPO1B  
 A:Note: sequence extracted from NCBI backbone (NCBIP:145988)  
 R:Gantz, I.; Shimoto, Y.; Konda, Y.; Dickinson, C.J.; Yamada, T.

Biochem. Biophys. Res. Commun. 200, 1214-1220, 1994  
A:Title: Molecular cloning, expression, and characterization of a fifth melanocortin receptor  
A:Reference number: JC2244; MUID:94241974  
A:Accession: JC2244

A:Molecule type: DNA  
A:Residues: 48-372 <GAN>

C:Cross-references: GB:L22527; NID:9468377; PIDN:AAA21337.1; PID:9468378  
C:Comment: This protein responds to melanocortins with an increase in intracellular cyclic AMP  
C:Superfamily: melanocortin receptor

C:Keywords: receptor; transmembrane protein

F:84-109/Domain: transmembrane #status predicted <TM1>

F:121-144/Domain: transmembrane #status predicted <TM2>

F:164-185/Domain: transmembrane #status predicted <TM3>

F:206-226/Domain: transmembrane #status predicted <TM4>

F:232-257/Domain: transmembrane #status predicted <TM5>

F:287-310/Domain: transmembrane #status predicted <TM6>

F:324-344/Domain: transmembrane #status predicted <TM7>

Query Match 14.0%; Score 282.5; DB 2; Length 372;  
Best Local Similarity 26.2%; Pred. No. 1.5e-14;  
Matches 85; Conservative 70; Mismatches 102; Indels 67; Gaps 14;

```
QY 33 GAGLRADAVVC-----LAVCAFI-----VLENALVLLGRHPRFHAPMFLLLGSLTSD 82
Db 70 GSNVKNKSLACEENGIAVEVFLTGLVSLLENLIVIGAIKKNKLNHSPMYFYVGS LAVAD 129

QY 83 LLAGAAYAANILLGSGPLTK-LSPALMFAR-----EGGVEVALTASVLSLAIALLERSL 135
Db 130 MLVSMNSNAWETVTIYLLNNKHLVADTFVPHIDNVFDSMICISVVASCSLAIADVRII 189

QY 136 TM-----ARRGPAPVSSRGRTLAWAAAAGVSLLLGLLPALGN-CIGRLDACS 183
Db 190 TIFVALRYHHIMTARRS-----GVIIACIWTFCI-----SCG 221

QY 184 TVLPYAKA-YVLPFCVLA-FVGILAAICAYRIYCOVRANARRLPARPGTAGTTSTRAR 241
Db 222 IVFIYYESKYVICLISMEFTMLFVWSLYTHMFLARNHVKRIAASP---RYNSVRQR 278

QY 242 RKPRSLALLRTLSVLLAFVACNGPLFLLLLLDVACPART-CPVLIQ-ADPFGLAMANS 299
Db 279 TSMKGAI---TLTLLGIFIVCWSPFHLILMISCPQNVYSCFMSFYNMVYLILIMONS 335

QY 300 LLNPIVITFNDRDLRHALLRLVCC 323
Db 336 VIDPLIYALRSQEMRRRTFKEIVCC 359
```

Search completed: January 16, 2002, 23:41:21  
Job time: 363 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 16, 2002, 23:38:53 ; Search time 64.41 Seconds  
(without alignments)  
226.558 Million cell updates/sec

Title: US-09-842-316-2

Perfect score: 2019

Sequence: 1 MESGLRPAPVSEVIVLHYN.....TGSPCAPTAARTLVSEPAAD 398

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 segs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_39.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	865.5	42.9	382	1	EDG1_MOUSE
2	853.5	42.3	383	1	EDG1_RAT
3	834	41.3	381	1	EDG1_HUMAN
4	774.5	38.4	378	1	EDG3_HUMAN
5	732.5	36.3	352	1	H218_RAT
6	518.5	25.7	364	1	EDG2_MOUSE
7	506.5	25.1	364	1	EDG2_HUMAN
8	505.5	25.0	393	1	EDG2_BOVIN
9	505.5	25.0	393	1	EDG2_SHEEP
10	440.5	21.8	180	1	EDG1_MOUSE
11	357	17.7	330	1	GPR3_HUMAN
12	352.5	17.5	334	1	GPR3_HUMAN
13	350	17.3	334	1	GPR3_MOUSE
14	347	17.2	330	1	GPR3_MOUSE
15	347	17.2	468	1	CB1A_FUGRU
16	346.5	17.2	334	1	GPR3_RAT
17	342	16.9	470	1	CB1B_FUGRU
18	325.5	16.1	473	1	CB1R_MOUSE
19	325.5	16.1	473	1	CB1R_RAT
20	323.5	16.0	472	1	CB1R_FELCA
21	322.5	16.0	473	1	CB1R_POEGU
22	318.5	15.8	473	1	CB1R_TARGR
23	318	15.8	362	1	GPR6_HUMAN
24	310	15.4	472	1	CB1R_HUMAN
25	308	15.3	363	1	GPR6_RAT
26	291.5	14.4	360	1	CB2R_HUMAN
27	289.5	14.3	294	1	MC5R_PIG
28	287.5	14.2	325	1	MC5R_BOVIN
29	283.5	14.0	325	1	MC5R_MOUSE
30	283	14.0	325	1	MC5R_MOUSE
31	282	14.0	466	1	BIAR_MOUSE
32	281.5	13.9	325	1	MC5R_SHEEP
33	.281	13.9	466	1	BIAR_RAT

34	277	13.7	325	1	MC5R_PANTR	O9tt23 pan troglod
35	276	13.7	473	1	BIAR_CANFA	P79148 canis famil
36	275	13.6	325	1	MC5R_RAT	P35345 rattus norv
37	275	13.6	332	1	MC4R_PIG	O97504 sus scrofa
38	273.5	13.5	347	1	CB2R_MOUSE	P47936 mus musculu
39	273	13.5	332	1	MC4R_BOVIN	O991j8 bos taurus
40	273	13.5	467	1	BIAR_BOVIN	O9tt96 bos taurus
41	272.5	13.5	360	1	CB2R_RAT	O9qzn9 rattus norv
42	272.5	13.5	474	1	BIAR_FELCA	O9tst6 felis silve
43	272	13.5	332	1	MC4R_HUMAN	P32245 homo sapien
44	269	13.3	360	1	MC3R_HUMAN	P41968 homo sapien
45	267	13.2	480	1	BIAR_MACMU	P47899 macaca mula

## ALIGNMENTS

RESULT 1	EDG1_MOUSE	STANDARD;	PRT;	382 AA.
AC	O08530;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUL-1998 (Rel. 36, Last annotation update)			
DE	PROBABLE G PROTEIN-COUPLED RECEPTOR EDG-1.			
GN	EDG1.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID:10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BA13/C; TISSUE=Liver;			
RX	MEDLINE-97369927; PubMed-9226368;			
RA	Liu C.H., Hia T.;			
RT	"The mouse gene for the inducible G-protein-coupled receptor edg-1.";			
RL	Genomics 43:15-24(1997).			
CC	-!- FUNCTION: THIS INDUCIBLE EPITHELIAL CELL G-PROTEIN-COUPLED RECEPTOR MAY BE INVOLVED IN THE PROCESSES THAT REGULATE THE DIFFERENTIATION OF ENDOTHELIAL CELLS (BY SIMILARITY).			
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.			
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES WITH HIGHEST LEVELS IN BRAIN, SPLEEN AND HEART. LOWER LEVELS FOUND IN LUNG, PLACENTA, MUSCLE, LIVER, UTERUS AND KIDNEY. VERY LOW LEVELS IN TESTIS AND THYMUS.			
CC	-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL; U40811; AAC53294.1; -			
DR	GCRdb; GCR_1661; -			
DR	MGI; MGI:1096355; Edg1.			
DR	InterPro; IPR000276; GPCR_Rhodpsn.			
DR	Pfam; PF000601; 7tm_1; 1.			
DR	PRINTS; PR00237; GPCR_Rhodopsn.			
DR	PRINTS; PR00642; EDG1ORPHAN.			
DR	PROSITE; PS00237; G-PROTEIN_RECEPTOR_F1_1; 1.			
DR	PROSITE; PS00262; G-PROTEIN_RECEPTOR_F1_2; 1.			
KW	G-protein coupled receptor; Transmembrane; Glycoprotein;			
FT	Phosphorylation; Lipoprotein; Palmitate.			
FT	DOMAIN 1..46 EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 47..71 1 (POTENTIAL).			
FT	DOMAIN 72..78 CYTOPLASMIC (POTENTIAL).			
FT	TRANSMEM 79..107 2 (POTENTIAL).			
FT	DOMAIN 108..121 EXTRACELLULAR (POTENTIAL).			
FT	TRANSMEM 122..140 3 (POTENTIAL).			
FT	DOMAIN 141..159 CYTOPLASMIC (POTENTIAL).			

```

FT TRANSMEM 160 185 4 (POTENTIAL).
FT DOMAIN 186 201 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 202 222 5 (POTENTIAL).
FT TRANSMEM 223 256 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 257 278 6 (POTENTIAL).
FT TRANSMEM 279 294 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 295 315 7 (POTENTIAL).
FT TRANSMEM 316 382 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 328 328 PALMITATE (BY SIMILARITY).
FT MOD_RES 353 353 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 382 AA; A0048993491C8957 CRC64;

Query Match 42.9%; Score 865.5; DB 1; Length 382;
Best Local Similarity 48.5%; Pred. No. 4.9e-50;
Matches 181; Conservative 66; Mismatches 105; Indels 21; Gaps 6;

QY 13 EVIVLHNYTGKLR-GARYOPGAGLRADAVCLAVCAFTVLENLAVLLVGRHPRHAPM 71
DB 23 DIIVRHNYTGKLNIG--EKDHGKILTSVVFILICCFILLENIFVLTITWTKKFRPM 80

QY 72 FLGSLTSLDLGAGAAAYANILLSGPLTKLSPALWFAREGGVFVALTSVLSLALAL 131
DB 81 YFIGNLALSDLLAGVAYTANILLSGATYTKLTPAQWFLREGSMFVALSASVLSLALAI 140

QY 132 ERSITMARPPAPVSSRGRTLMAAAANGVSLLLGLPALGNCLGRDAGSTVLPYAK 191
DB 141 ERYITMLKMLKLNHNGSSRSFLLISACWISLILGLPSMGWNCISLSSCSTVLPYHK 200

QY 192 AYVLCVLAFLVGLIAATCALYARIYCOVRANARLPARPAGTAGTSTRARKPRSLALLR 251
DB 201 HYILFCVTFTLLLSLAILCYIYSLVTRSRRLTFK-----NISKGRSEKSLALLK 256

QY 252 TLSVLLAFVACGPFLLLLLDVACPARTCPVLLQADPFGLGLAMANSLLNPIIYTLNR 311
DB 257 TVIIVSVFACWAPFILLLDVGCARKTCDILYKAELVFLAVLNSGTPPIIYTLNK 316

QY 312 DLRALRLVCCGRHSGRDPSCSQSQAASAGLRCPLPPGLDGFSGSERSS--POR 370
DB 317 EMRAFAIRIV-----SCCKCPNG-----DSACKFKRPIIPGNEFSRSKSDNSHPQK 363

QY 371 DGLDTSSTGSPG 383
DB 364 DGDNPETIMSSG 376

RESULT 2
EDG1_RAT
ID EDG1_RAT STANDARD; PRT; 383 AA.
AC P48303;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PROBABLE G PROTEIN-COUPLED RECEPTOR EDG-1.
GN EDG1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RX MEDLINE=95047498; PubMed=7959012;
RA Lado D.C., Browe C.S., Gaskin A.A., Borden J.M., MacLennan A.J.;
RT "Cloning of the rat edg-1 immediate-early gene: expression pattern
RL suggests diverse functions.";
RL Gene 149:331-336(1994).
CC -1- FUNCTION: IT IS POSSIBLE THAT IT ACTS AS A RECEPTOR FOR A TROPIC
CC FACTOR AND IS INVOLVED IN THE SURVIVAL OF BRAIN CELLS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT EMBRYONIC DAY 15. AT
CC POSTNATAL DAY 14 DETECTED IN SKIN, SPLEEN, LIVER, KIDNEY, HEART,

```

```

CC TESTICLE, LUNG AND BRAIN. AT ADULTHOOD IS MOST ABUNDANT IN BRAIN.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U10303; AAA83418.1;
CC GCRdb; GCR_1453;
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCR_RHODOPSIN.
CC PRINTS; PR00642; EDG1ORPHANR.
CC PROSITE; PS00237; G-PROTEIN_RECP_F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN_RECP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Phosphorylation; Lipoprotein; Palmitate.
CC DOMAIN 1 47 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 48 72 1 (POTENTIAL).
CC DOMAIN 73 79 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 80 108 2 (POTENTIAL).
CC DOMAIN 109 122 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 123 141 3 (POTENTIAL).
CC DOMAIN 142 160 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 161 186 4 (POTENTIAL).
CC DOMAIN 187 202 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 203 223 5 (POTENTIAL).
CC DOMAIN 224 257 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 258 279 6 (POTENTIAL).
CC DOMAIN 280 295 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 296 316 7 (POTENTIAL).
CC DOMAIN 317 383 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 31 31 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC LIPID 329 329 PALMITATE (BY SIMILARITY).
CC MOD_RES 354 354 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 383 AA; 42745 MW; 090BA6AE09DB4F3 CRC64;

Query Match 42.3%; Score 853.5; DB 1; Length 383;
Best Local Similarity 47.6%; Pred. No. 3e-49;
Matches 177; Conservative 66; Mismatches 110; Indels 19; Gaps 5;

QY 13 EVIVLHNYTGKLRGARYOPGAGLRADAVCLAVCAFTVLENLAVLLVGRHPRHAPMF 72
DB 24 DIIVRHNYTGKL-NIGVEKDHGKILTSVVFILICCLILENIFVLTITWTKKFRPMY 82

QY 73 LILGSLTSLDLGAGAAAYANILLSGPLTKLSPALWFAREGGVFVALTSVLSLALALE 132
DB 83 YFIGNLALSDLLAGVAYTANILLSGATYTKLTPAQWFLREGSMFVALSASVLSLALAI 142

QY 133 RSLTMRARPAPVSSRGRTLMAAAANGVSLLLGLPALGNCLGRDAGSTVLPYAKA 192
DB 143 RYITMLKMLKLNHNGSSRSFLLISACWISLILGLPLMGWNCISLSSCSTVLPYHKH 202

QY 193 VYLCVLAFLVGLIAATCALYARIYCOVRANARLPARPAGTAGTSTRARKPRSLALLR 252
DB 203 YILFCVTFTLLLSLAILCYIYSLVTRSRRLTFK-----NISKGRSEKSLALLK 258

QY 253 LSVLLAFVACGPFLLLLLDVACPARTCPVLLQADPFGLGLAMANSLLNPIIYTLNRD 312
DB 259 VIIVLSVFIACWAPFILLLDVGCARKTCDILYKAELVFLAVLNSGTPPIIYTLNKE 318

QY 313 LEHALRLVCCGRHSGRDPSCSQSQAASAGLRCPLPPGLDGFSGSERSS--PORD 371
DB 319 MFAIRIRII-----SCCKCPNG-----DSACKFKRPIIPGNEFSRSKSDNSHPQK 365

QY 372 GLDTSSTGSPG 383
DB 366 DGDNPETIMSSG 377

```









Db 79 RFHPIYLLMANLAAADFFAGLAYFYLMENTGPNTRRLTVSTWLLRQGLDTSLTASVAN 138  
 QY 126 LLAIALERSLUTWARGPAPVSSRGHTLMAAAANGVSLLLGLLPALGWNCLGRDACSTV 185  
 Db 139 LLAIAERHITVFMOLHTRMSNRVVVVIWTHAIVMGAIPSVGVNCCIDHCNSNM 198  
 QY 186 LPLXAKAYVLCVAFVGLAALCALYARIYCOVRANARLPARGGTAGTSTRARRKPR 245  
 Db 199 APLYSDSLVAFWAIFNLVTFVVMVLYAHFGYVQRTMRSR-----SSGPRNRDT 252  
 QY 246 SLALLRTSLVLLAFVACWGLFLLLLDVACPARTCPVLLQADPFGLAMANSLLNPPI 305  
 Db 253 MMSLLKTVVILGAFIVCTGLVLLLDVCCP--QCDVLAYEKFFLLAEFNSANPPI 310  
 QY 306 YLTNRDLRHALLRLVCCGRHSCGRDPS--GSOQSASA 341  
 Db 311 YSYRDKEMSATFRQILCCORNENGTGESSDRSASS 347

RESULT 7  
 EDG2\_HUMAN STANDARD; PRT; 364 AA.  
 AC Q92633; O00656; P78351;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE LYOPHOSPHATIDIC ACID RECEPTOR (EDG-2).  
 GN EDG2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_Taxid-9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-97224397; PubMed-9070858;  
 RA Moolenaar W.H., Kranenburg O., Postma F.R., Zondag G.C.M.;  
 RT "Lysophosphatidic acid: G-protein signalling and cellular responses";  
 RL Curr. Opin. Cell Biol. 9:168-173(1997).  
 CC -1- FUNCTION: RECEPTOR FOR LYOPHOSPHATIDIC ACID (LPA), A MEDIATOR OF  
 DIVERSE CELLULAR ACTIVITIES.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.ebi.ac.uk/announcements> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; U08011; AAC51139.1; -  
 CC EMBL; Y09479; CAAY0686.1; -  
 CC EMBL; Y09479; CAAY0687.1; -  
 CC EMBL; U78192; AAC00530.1; -  
 CC GCRDB; GCR\_1342; -  
 CC GCRDB; GCR\_1897; -  
 CC GCRDB; GCR\_1898; -  
 CC MIN; 602282; -  
 CC InterPro; IPR000276; GPCR\_Rhodpsn.  
 CC Pfam; PF00001; 7tm\_1; 1.  
 CC PRINTS; PR01148; LPARECEPTOR.  
 CC PROSITE; PS00237; G-PROTEIN\_RECP\_F1\_1; 1.  
 CC PROSITE; PS50262; G-PROTEIN\_RECP\_F1\_2; 1.  
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Phosphorylation; Lipoprotein; Palmitate.

FT DOMAIN 1 50  
 FT TRANSSEM 51 75  
 FT DOMAIN 76 82  
 FT TRANSSEM 83 111  
 FT DOMAIN 112 125  
 FT TRANSSEM 126 144  
 FT DOMAIN 145 163  
 FT TRANSSEM 164 189  
 FT DOMAIN 190 205  
 FT TRANSSEM 206 226  
 FT DOMAIN 227 258  
 FT TRANSSEM 259 280  
 FT DOMAIN 281 294  
 FT TRANSSEM 295 315  
 FT DOMAIN 316 364  
 FT LIPID 327 327  
 FT CARBOHYD 35 35  
 FT CARBOHYD 35 35  
 FT CONFLICT 340 340  
 SQ SEQUENCE 364 AA; 41139 MW; 59E772FD00DFE74 CRC64;

Query Match 25.1%; Score 506.5; DB 1; Length 364;  
 Best Local Similarity 32.9%; Pred. No. 1.4e-26;  
 Matches 112; Conservative 73; Mismatches 130; Indels 25; Gaps 5;

QY 12 SEVIVLHYNTGKRGARYQP-----GAGLRADAVYVCAVFIENLAVLLVLRHP 65  
 Db 27 NESIAFPIYNNRSGRHLATEWNTVSKVLGLG-----ITVCIFIMLANLLVMAIYVNR 78  
 QY 66 EFHAPFLILGSLTDLLAGAAYANILLGSLTLKLSPALWFAREGGVFAVLTASVLS 125  
 Db 79 EFHPIYLLMANLAAADFFAGLAYFYLMENTGPNTRRLTVSTWLLRQGLDTSLTASVAN 138  
 QY 126 LLAIALERSLUTWARGPAPVSSRGHTLMAAAANGVSLLLGLLPALGWNCLGRDACSTV 185  
 Db 139 LLAIAERHITVFMOLHTRMSNRVVVVIWTHAIVMGAIPSVGVNCCIDHCNSNM 198  
 QY 186 LPLXAKAYVLCVAFVGLAALCALYARIYCOVRANARLPARGGTAGTSTRARRKPR 245  
 Db 199 APLYSDSLVAFWAIFNLVTFVVMVLYAHFGYVQRTMRSR-----SSGPRNRDT 252  
 QY 246 SLALLRTSLVLLAFVACWGLFLLLLDVACPARTCPVLLQADPFGLAMANSLLNPPI 305  
 Db 253 MMSLLKTVVILGAFIVCTGLVLLLDVCCP--QCDVLAYEKFFLLAEFNSANPPI 310  
 QY 306 YLTNRDLRHALLRLVCCGRHSCGRDPS--GSOQSASA 341  
 Db 311 YSYRDKEMSATFRQILCCOR---SENPTGPTESSDRSASS 347

RESULT 8  
 EDG2\_BOVIN STANDARD; PRT; 364 AA.  
 ID EDG2\_BOVIN  
 AC Q28031;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE LYOPHOSPHATIDIC ACID RECEPTOR (EDG-2) (RECI.3).  
 GN EDG2.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_Taxid-9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE-97165887; PubMed-9013780;  
 RA Macrae A.D., Premont R.T., Jaber M., Petersen A.S., Lefkowitz R.J.;  
 RT "Cloning, characterization, and chromosomal localization of recl.3, a  
 RT member of the G-protein-coupled receptor family highly expressed in  
 RT brain.";

Brain Res. Mol. Brain Res. 42:245-254(1996).

-1- FUNCTION: RECEPTOR FOR LYSOPHOSPHATIDIC ACID (LPA), A MEDIATOR OF DIVERSE CELLULAR ACTIVITIES.

-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U48236; AAC48695.1; -

CC DB; GCR\_1218; -

DR InterPro: IPR000276; GPCR\_Rhodpsn.

DR Pfam: PF00001; 7tm1; 1.

DR PRINTS: PRO1148; LPARECEPTOR.

DR PROSITE: PS00237; G-PROTEIN\_RECF\_F1\_1; 1.

DR PROSITE: PS00262; G-PROTEIN\_RECF\_F1\_2; 1.

DR G-protein coupled receptor; Transmembrane; Glycoprotein;

KW Phosphorylation; Lipoprotein; Palmitate.

KW DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 51 75 1 (POTENTIAL).

FT DOMAIN 76 82 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 83 111 2 (POTENTIAL).

FT DOMAIN 112 125 3 (POTENTIAL).

FT TRANSMEM 126 144 4 (POTENTIAL).

FT DOMAIN 145 163 5 (POTENTIAL).

FT TRANSMEM 164 189 6 (POTENTIAL).

FT DOMAIN 190 205 7 (POTENTIAL).

FT TRANSMEM 206 226 8 (POTENTIAL).

FT DOMAIN 227 258 9 (POTENTIAL).

FT TRANSMEM 259 280 10 (POTENTIAL).

FT DOMAIN 281 294 11 (POTENTIAL).

FT TRANSMEM 295 315 12 (POTENTIAL).

FT DOMAIN 316 336 13 (POTENTIAL).

FT TRANSMEM 337 357 14 (POTENTIAL).

FT LIPID 357 357 PALMITATE (BY SIMILARITY).

FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).

SQ SEQUENCE 364 AA; 41070 MW; DE94675BA407B2F6 CRC64;

Query Match 25.0%; Score 505.5; DB 1; Length 364;

Best Local Similarity 33.5%; Pred. No. 1.6e-26;

Matches 113; Conservative 70; Mismatches 131; Indels 23; Gaps 5;

QY 12 SEVIVLHYNTGKLGARYQP-----GAGLRADAVVCLAVCAFTVLENLAVLLVLRHP 65

DB 27 NESIAFFYNSRGKYLATENTVTYKLVMLG-----ITVCIFIMLANLLVWVAIYVNR 78

QY 66 RFIAPFELLGSLTSLDLAAGAAAYANILLSPGLKLSPALFWAREGVFVALTSVLS 125

DB 79 RFHPIYILMANLAADFFAGLAYFYLMFNTGTRRLTSTVTLRQGLIDTSLTVSVAN 138

QY 126 LLAIALERSLTMARPGAPVSSRGRTLMAAAAGVSLGLLGLPALGNCLGRDLDCSTV 185

DB 139 LLAIAIERHTVPRQLHARMSNRVVVVIVTWMAIVMGAIPIVSGNCTCDIENCSNM 198

QY 186 LPIYAKAYVLCVLAFLVLAICALYARIYCOVRRANRRLPARGTGTSTRARRKPR 245

DB 199 APLYSDSYLVFAIFNLVTFVVMVLYAHIFGYVQRTRMSRH-----SSGPRNRDT 252

QY 246 SLALLRTSLVLLFAVACGFLPLLLLDVACPARTCPVLLQADPFLGLMANSLNPII 305

DB 253 MMSLLTIVVLGAFICFTICPGVLLLDVCCP--QCDVLAKEFLLAIFNSANPII 310

QY 306 YTLNRLRLLRLLVCCGRHSCGRDPS--GSOQSASA 341

DB 311 YSYRDKMSATFQILCCQRSENSTSGPTSGDRSASS 347

RESULT 9

EDG2\_SHEEP

ID EDG2\_SHEEP STANDARD; PRT; 393 AA.

AC P46628;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE LYSOPHOSPHATIDIC ACID RECEPTOR (EDG-2).

GN EDG2 OR EDG-2.

OS Ovis aries (Sheep).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Caprinae; Ovis.

OX NCBI\_TaxID:9940;

RN [1]

RP SEQUENCE FROM N.A.

RX TISSUE=Follicular pars tuberalis;

RX MEDLINE=96430916; PubMed=8833998;

RA Masana M.I., Brown R.C., Pu H., Gurney M.E., Dubocovich M.L.;

RT "Cloning and characterization of a new member of the G-protein coupled receptor EDG family."

RL Recept. Channels 3:255-262(1995).

CC -1- FUNCTION: RECEPTOR FOR LYSOPHOSPHATIDIC ACID (LPA), A MEDIATOR OF DIVERSE CELLULAR ACTIVITIES.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; U18405; AAB52368.1; -

CC GCRDB; GCR\_1271; -

DR InterPro: IPR000276; GPCR\_Rhodpsn.

DR Pfam: PF00001; 7tm1; 1.

DR PRINTS: PRO1148; LPARECEPTOR.

DR PROSITE: PS00237; G-PROTEIN\_RECF\_F1\_1; 1.

DR PROSITE: PS00262; G-PROTEIN\_RECF\_F1\_2; 1.

DR G-protein coupled receptor; Transmembrane; Glycoprotein;

KW Phosphorylation; Lipoprotein; Palmitate.

FT DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 51 75 1 (POTENTIAL).

FT DOMAIN 76 82 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 83 111 2 (POTENTIAL).

FT DOMAIN 112 125 3 (POTENTIAL).

FT TRANSMEM 126 144 4 (POTENTIAL).

FT DOMAIN 145 163 5 (POTENTIAL).

FT TRANSMEM 164 189 6 (POTENTIAL).

FT DOMAIN 190 205 7 (POTENTIAL).

FT TRANSMEM 206 226 8 (POTENTIAL).

FT DOMAIN 227 258 9 (POTENTIAL).

FT TRANSMEM 259 280 10 (POTENTIAL).

FT DOMAIN 281 294 11 (POTENTIAL).

FT TRANSMEM 295 315 12 (POTENTIAL).

FT DOMAIN 316 336 13 (POTENTIAL).

FT LIPID 337 327 PALMITATE (BY SIMILARITY).

FT CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 35 35 N-LINKED (GLCNAC. .) (POTENTIAL).

SQ SEQUENCE 393 AA; 44433 MW; 356E961153C345FC CRC64;

Query Match 25.0%; Score 505.5; DB 1; Length 393;

Best Local Similarity 33.5%; Pred. No. 1.7e-26;

Matches 113; Conservative 70; Mismatches 131; Indels 23; Gaps 5;

QY 12 SEVIVLHYNTGKLGARYQP-----GAGLRADAVVCLAVCAFTVLENLAVLLVLRHP 65

DB 27 NESIAFFYNSRGKYLATENTVTYKLVMLG-----ITVCIFIMLANLLVWVAIYVNR 78

QY 66 RHAPMELLGSLTSLDLAGAAYANILLSGPLTLKLSPALWFAREGGVVALTASVLS 125  
 DB 79 REFFIITLMANLAADFAAGLAFYLMFTGNTRTLVSTWLLROGLDITVTASVAN 138  
 QY 126 LLAIALERSLTMRGAPVSSRGRTLAMAAAAGVSLIGLPLGALGNCLGRLDACSTV 185  
 DB 139 LLAIAIERHITVERMOLTRMSNRVVVVIWIAVNGAIPSGVNCICDIENCSNM 198  
 QY 186 LPLAKAYVLCVLAFCVLAICALYARIYCOVRANARLARPCTAGTTSTRARRKPR 245  
 DB 199 APLYSDSYLVFWALFNLFVFMVVLVAHIFGVVORTMMSRH-----SSGPRNRDT 252  
 QY 246 SLALRTLSVLLAFVACWGLFLLLLDVACPARTCPVLLQADPFLGLAMNSLNPIT 305  
 DB 253 WMSLLKTVVILGAFICWFLGVLVLLLDVCCP--QCDVLAYEKFLLAEFNSAMNPIT 310  
 QY 306 YLTNRDRHALLRLVCGRHSGRDP--GSOOSASA 341  
 DB 311 YSRDKENSATFROILCORSENTSGTSGDRSASS 347  
 RESULT 10  
 EDGL\_MOUSE STANDARD; PRT; 180 AA.  
 AC P52592;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE PROBABLE G PROTEIN-COUPLED RECEPTOR EDG-1 LIKE (FRAGMENT).  
 GN GPCR13.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Testis;  
 RX MEDLINE=94116980; PubMed=8288218;  
 RA Wilkie T.M., Chen Y., Gilbert D.J., Moore K.J., Yu L., Simon M.I.,  
 RA Copeland N.G., Jenkins N.A.;  
 RT "Identification, chromosomal location, and genome organization of  
 RT mammalian G-protein-coupled receptors."  
 RL Genomics 18:175-184(1993).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L20334; AAA16846.1; -  
 DR GCRDB; GCR\_0811; -  
 DR MGD; MGI:95569; Gpcr13  
 DR InterPro; IPR000226; GPCR\_Rhodopsn.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEPTOR\_F1\_1; PARTIAL.  
 DR PROSITE; PS0262; G-PROTEIN\_RECEPTOR\_F1\_2; 1.  
 KW G-protein coupled receptor; Transmembrane.  
 FT NON\_TER 1 1  
 FT DOMAIN <1 5  
 FT TRANSNEM 6 34  
 FT DOMAIN 35 48  
 FT TRANSNEM 49 67  
 FT DOMAIN 68 86  
 FT TRANSNEM 87 112  
 FT DOMAIN 113 128  
 FT TRANSNEM 129 149  
 FT DOMAIN 150 172  
 FT TRANSNEM 173 >180  
 FT CYTOPLASMIC (POTENTIAL).  
 FT 2 (POTENTIAL).  
 FT EXTRACELLULAR (POTENTIAL).  
 FT 3 (POTENTIAL).  
 FT CYTOPLASMIC (POTENTIAL).  
 FT 4 (POTENTIAL).  
 FT EXTRACELLULAR (POTENTIAL).  
 FT 5 (POTENTIAL).  
 FT CYTOPLASMIC (POTENTIAL).  
 FT 6 (POTENTIAL).

FT NON\_TER 180 180  
 SQ SEQUENCE 180 AA; 19364 MW; 91D99CF8769C4CBD CRC64;  
 Query Match 21.8%; Score 440.5; DB 1; Length 180;  
 Best Local Similarity 50.5%; Pred. No. 1.5e-22;  
 Matches 97; Conservative 27; Mismatches 53; Indels 15; Gaps 2;  
 QY 66 RHAPMELLGSLTSLDLAGAAYANILLSGPLTLKLSPALWFAREGGVVALTASVLS 125  
 DB 2 KEHSAMVLEGLNLAASDLLAGVAFVANTLLSGHVTLSTLPVQVFAREVSFAITLSASVFS 61  
 QY 126 LLAIALERSLTMRGAPVSSRGRTLAMAAAAGVSLIGLPLGALGNCLGRLDACSTV 185  
 DB 62 LLAIAIERHITVERMOLTRMSNRVVVVIWIAVNGAIPSGVNCICDIENCSNM 121  
 QY 186 LPLAKAYVLCVLAFCVLAICALYARIYCOVRANARLARPCTAGTTSTRARRKPR 245  
 DB 122 LPLAKRYVLCVVTIESVILLAILAIVALYVRIYFVRS-----SHADVAG-----PQ 166  
 QY 246 SLALRTLSVVL 257  
 DB 167 TLALKTVITVL 178  
 RESULT 11  
 GPR3\_HUMAN STANDARD; PRT; 330 AA.  
 AC P46089;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PROBABLE G PROTEIN-COUPLED RECEPTOR GPR3 (ACCA ORPHAN RECEPTOR).  
 GN GPR3 OR ACCA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96015070; PubMed=8530049;  
 RA Song Z.-H., Modi W., Bonner T.I.;  
 RT "Molecular cloning and chromosomal localization of human genes  
 RT encoding three closely related G protein-coupled receptors."  
 RL Genomics 28:347-349(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95213036; PubMed=7698767;  
 RA Iismaa T.P., Kiefer J., Liu M.L., Baker E., Sutherland G.R.,  
 RA Shine J.;  
 RT "Isolation and chromosomal localization of a novel human G-protein-  
 RT coupled receptor (GPR3) expressed predominantly in the central  
 RT nervous system."  
 RL Genomics 24:391-394(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95366960; PubMed=7639700;  
 RA Eggerickx D., Deneef J.F., Labbe O., Hayashi Y., Refetoff S.,  
 RA Vassart G., Parmentier M., Libert F.;  
 RT "Molecular cloning of an orphan G-protein-coupled receptor that  
 RT constitutively activates adenylylate cyclase."  
 RL Biochem. J. 309:837-843(1995).  
 RN [4]  
 RP SEQUENCE OF 1-292 FROM N.A.  
 RX MEDLINE=95154831; PubMed=7851889;  
 RA Marchese A., Docherty J.M., Nguyen T., Helber M., Cheng R.,  
 RA Heng H.H.O., Tsui L.-C., Shi X., George S.R., O'Dowd B.F.;  
 RT "Cloning of human genes encoding novel G protein-coupled receptors."  
 RL Genomics 23:609-618(1994).  
 CC -1- FUNCTION: ORPHAN RECEPTOR. THE ACTIVITY OF THIS RECEPTOR IS  
 CC MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYLATE CYCLASE.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE CENTRAL NERVOUS



Query Match	17.5%;	Score	352.5;	DB	1;	Length	334;
Best Local Similarity	34.6%;	Pred. No.	1.5e-16;				
Matches	101;	Conservative	44;	Mismatches	120;	Indels	27;
Gaps							
10:							
QY	39	DAYVCVACAFIVLENLAVLLVLRGHRPRFHAPMFLLIGSLTSLDLAGAAYAAANTILLGP	98				
Db	48	DIVLCTS-CTLISCENAIWLLIFHNPSLRAPMFLIGSLADLAGLIGLITNEVFA--	104				
QY	99	LTKLSPALWFAREGGVFVALTASVLLIAIAIERSLT----	NARGPAPVSGRGETLAM	154			
Db	105	-YLQSEATKVTIGLIVASFSASVCSLLHATVDRIYLSLYATYHSETVTTF--	TYVM	160			
QY	155	AAAMGVSLLLGLLPAWGNCIGRLDACSTVLPYAKAYVLCVLAFAVGILAAICAYAR	214				
Db	161	LVMLAGTSTICGLGFLVPMGNCRLDESTCSVVRPL-TKNNAAILSVSFLFMFLMLQLYIQ	219				
QY	215	IYQVRANARRLPAPRGTAGTTSTRARKPRSLALLRTLSVLLAFVACWGPFLFL-LLL	273				
Db	220	ICKIVMRHAHQALQHHELAHSHTVYTRKGVS-----	TLAILGTFACWMPFTYLSLIA	274			
QY	274	DVACPA-RTCPVVLQADPFLGLAMANSLNPIIYTTNDRDLRALRLVCCG	324				
Db	275	DITYFSIYTAFLP-----	ATYNSIINPIYAFNBOIQKALC-LIICC	318			

```

RESULT 13
GPRC_MOUSE
ID GPRC_MOUSE STANDARD; PRT; 334 AA.
AC P35412;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE PROBABLE G PROTEIN-COUPLED RECEPTOR GPR12 (GPCR01).
GN GPR12 OR GPCR12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OC NCBI_Taxid=10090;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=BALB/C; TISSUE=Brain;
RX MEDLINE=94085630; PubMed=8262253;
RA Saeiki Y., Ueno S., Mizuno R., Nishimura T., Fujimura H., Negai Y.,
RA Yanggihara T.;
RA "Molecular cloning of a novel putative G protein-coupled receptor
RT (GPCR21) which is expressed predominantly in mouse central nervous
RT system.;"
RL FEBS Lett. 336:317-322(1993).
RL -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -1- FUNCTION: ORPHAN RECEPTOR.
CC -1- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE FOREBRAIN AND A
CC LESSER EXTENT IN THE HINDBRAIN. LOWER EXPRESSION IN THE LIVER.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL; D21061; BAA04640.1; -
CC GCRDB; GCR_0902; -
CC MGI; MGI:101909; Gpcr12.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR005644; GPCRORPHANR.
CC PRINTS; PR00550; GPR12ORPHANR.
CC PROSITE; PS00237; G-PROTEIN_RECEPT_F1_1; 1.
CC PROSITE; PS00532; G-PROTEIN_RECEPT_F1_2; 1.
CC -----

```













XX	Human PFI-006 encoding nucleotide sequence SEQ ID NO:1.
DE	
XX	
KW	Human; PFI-006; G-protein coupled receptor; GPCR; chromosome 19;
KW	anorectic; antidiabetic; neuroprotective; neuroleptic; nootropic;
KW	psychotropic; antiinflammatory; cytostatic; vulnary; dermatological;
KW	osteopathic; cardiovascular; antiallergic; gene therapy; diabetes;
KW	G-protein coupled receptor modulator; signal transduction; obesity;
KW	metabolic disease; neurological disease; psychotherapeutic; inflammation;
KW	urogenital disease; reproduction; sexual medicine; cancer; tissue repair;
KW	dermatology; skin pigmentation; photoaging; frailty; osteoporosis; ds;
KW	cardiovascular disease; gastrointestinal disease; antinfection; allergy;
KW	respiratory disease; sensory organ disorder; sleep disorder; hair loss.
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 1..1198
FT	/tag= a
FT	/product= "PFI-006"
FT	/note= "putative G-protein coupled receptor"
XX	
PN	EPI090925-Al.
XX	
PD	11-APR-2001:
XX	
PF	06-OCT-2000; 2000EP-0308853.
XX	
PR	08-OCT-1999; 99GB-0023890.
XX	
PA	(PFIZ ) PFIZER LTD.
PA	(PFIZ ) PFIZER INC.
XX	
PI	Harland L;
XX	
WI	WPI: 2001-302043/32.
DR	P-PSDB; AAB74956.
XX	
PT	New human G-protein coupled receptor polynucleotide and polypeptide,
PT	useful in drug screening, as well as for diagnosing or treating
PT	diseases associated with signal transduction, e.g. obesity,
PT	inflammation, tissue repair or cancer .
XX	
PS	Claim 1; Page 36; 46pp; English.
XX	
CC	The present sequence encodes human PFI-006 which is a putative G-protein
CC	coupled receptor (GPCR). PFI-006 was identified from chromosome 19.
CC	PFI-006 has anorectic, antidiabetic, neuroprotective, neuroleptic,
CC	nootropic, psychotropic, antiinflammatory cytotactic, vulnary,
CC	dermatological, osteopathic, cardiovascular and anti allergic activities.
CC	It can be used in pharmaceuticals or in the manufacture of medicaments
CC	for the treatment of a patient having need to modulate, i.e. agonise or
CC	(selectively) antagonise PFI-006 protein. A compound which modulates
CC	PFI-006 can be used in gene therapy. PFI-006 polynucleotides and
CC	proteins can be used in diagnosing and treating a variety of diseases.
CC	These are also useful for evaluating and/or screening agents that can
CC	modulate the GPCR. In particular, they are useful for screening drug
CC	candidates for treating diseases associated with signal transduction,
CC	including obesity, diabetes and metabolic disease, neurological disease,
CC	psychotherapeutics, urogenital disease, reproduction and sexual
CC	medicine, inflammation, cancer, tissue repair, dermatology, skin
CC	pigmentation, photoaging, frailty, osteoporosis, cardiovascular disease,
CC	gastrointestinal disease, antinfection, allergy and respiratory
CC	disease, sensory organ disorders, sleep disorders or hair loss.
XX	
SQ	Sequence 1198 BP; 139 A; 449 C; 389 G; 221 T; 0 other;
XX	
Query Match	99.8%; Score 1195; DB 22; Length 1198;
Best Local Similarity	100.0%; Pred. No. 5.3e-216;
Matches 1195; Conservative	0; Mismatches 0; Indels 0; Gaps 0
Y	i atggagtcggggctgtcgcggccgcccgatgcagcgaagtcatcgtctcgtattacaac 60



QY 901 ctgaaccatcatcatcagctcaccacacacgagcctgcccacgagcctcctcgcctg 960  
Db 1052 ctgaaccatcatcatcagctcaccacacgagcctgcccacgagcctcctcgcctg 1111  
QY 961 gtctgtcggagcagcactcctcgcgagagacacgagctgctccacagcagtcggcagc 1020  
Db 1112 gtctgtcggagcagcactcctcgcgagagacacgagctgctccacagcagtcggcagc 1171  
QY 1021 gctgtcgtcggagcagcactcctcgcgagagacacgagctgctccacagcagtcggcagc 1080  
Db 1172 gctgtcgtcggagcagcactcctcgcgagagacacgagctgctccacagcagtcggcagc 1231  
QY 1081 agcggctcggagcagcactcctcgcgagagacacgagctgctccacagcagtcggcagc 1140  
Db 1232 agcggctcggagcagcactcctcgcgagagacacgagctgctccacagcagtcggcagc 1291  
QY 1141 agcggctcggagcagcactcctcgcgagagacacgagctgctccacagcagtcggcagc 1197  
Db 1292 agcggctcggagcagcactcctcgcgagagacacgagctgctccacagcagtcggcagc 1348  
RESULT 5  
AAD01132  
ID AAD01132 standard; cDNA; 1503 BP.  
XX  
AC AAD01132;  
XX  
DT 02-NOV-2000 (first entry)  
XX  
DE Human orphan G protein-coupled receptor hCHN6 cDNA.  
XX  
KW Human; orphan G protein-coupled receptor; GPCR; hCHN6; drug screening;  
KW transmembrane receptor; expressed sequence tag; EST; signal cascade; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1503  
FT /tag= a  
FT /product= "hCHN6"  
FT /note= "Human orphan G protein-coupled receptor"  
XX  
PN WO200031258-A2.  
XX  
PD 02-JUN-2000.  
XX  
PF 13-OCT-1999; 99WO-US23687.  
XX  
PR 20-NOV-1998; 98US-0109213.  
PR 16-FEB-1999; 98US-0120416.  
PR 26-FEB-1999; 98US-0121852.  
PR 12-MAR-1999; 98US-0123946.  
PR 12-MAR-1999; 98US-0123949.  
PR 28-MAY-1999; 98US-0136436.  
PR 28-MAY-1999; 98US-0136437.  
PR 28-MAY-1999; 98US-0136439.  
PR 28-MAY-1999; 98US-0136567.  
PR 28-MAY-1999; 98US-0137127.  
PR 28-MAY-1999; 98US-0137131.  
PR 29-JUN-1999; 98US-0141448.  
PR 29-SEP-1999; 98US-0156555.  
PR 29-SEP-1999; 98US-0156633.  
PR 29-SEP-1999; 98US-0156634.  
PR 29-SEP-1999; 98US-0156653.  
PR 01-OCT-1999; 98US-0157280.  
PR 01-OCT-1999; 98US-0157281.  
PR 01-OCT-1999; 98US-0157282.  
PR 01-OCT-1999; 98US-0157293.  
PR 01-OCT-1999; 98US-0157294.  
PR 12-OCT-1999; 98US-0416760.  
PR 12-OCT-1999; 98US-0417044.  
XX  
PA (AREN-) ARENA PHARM INC.

XX  
PI Chen R, Dang HT, Liaw CW, Lin I;  
XX WPI; 2000-400068/34.  
DR P-PSDB; AAY71305.  
XX  
PT Novel human orphan G protein-coupled receptors and the encoding cDNAs  
PT for use in the identification of G protein-coupled receptor agonists -  
PS Claim 57; Page 78-79; 102pp; English.  
XX  
CC The present sequence is a cDNA encoding hCHN6, an endogenous human  
CC orphan G protein-coupled receptor (GPCR). The hCHN6 cDNA was identified  
CC using ESTs (expressed sequence tag) AA804531 and 2134670 as a probe.  
CC The orphan GPCR of the invention, like all GPCRs has seven transmembrane  
CC alpha helices with an extracellular N-terminus and an intracellular  
CC C-terminus. However, no endogenous ligands has yet been identified for  
CC the proteins of the invention. The orphan GPCRs may be used in the  
CC identification of their endogenous ligands, and to screen potential GPCR  
CC agonists and antagonists for use as pharmaceutical agents. The proteins  
CC may also be used in the study of GPCR-mediated signalling cascades, and  
CC to elucidate their precise role in normal and diseased human conditions.  
CC Nucleic acid encoding human orphan GPCRs may be used for tissue  
CC localisation expression analysis to provide information about their  
CC function in healthy and pathological states.  
XX  
SQ Sequence 1503 BP; 176 A; 562 C; 511 G; 254 T; 0 other;

Query Match 99.6%; Score 1192.2; DB 21; Length 1503;  
Best Local Similarity 99.7%; Pred. No. 1.8e-215;  
Matches 1194; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 atggagtcggggctgctgagcggcgccggcgagcagagtcctcctgcattacac 60  
Db 307 atggagtcggggctgctgagcggcgccggcgagcagagtcctcctgcattacac 366  
QY 61 tacaccggcaagctcccggtgctgctaccagccgggtgcccggcgcgcgcagcgc 120  
Db 367 tacaccggcaagctcccggtgctgctaccagccgggtgcccggcgcgcgcagcgc 426  
QY 121 ggggtgctcggcggtgctgagccttcctgctgctagagaaatctagcgtgttgggtg 180  
Db 427 ggggtgctcggcggtgctgagccttcctgctgctagagaaatctagcgtgttgggtg 486  
QY 181 ctggagccaccgcgcttcacagctccacgtgttcctgctcctggcagcctcacgttg 240  
Db 487 ctggagccaccgcgcttcacagctccacgtgttcctgctcctggcagcctcacgttg 546  
QY 241 tgggactgctggcagcgccgctacgcgcgcaacacatcctactgtggggccgctcacg 300  
Db 547 tgggactgctggcagcgccgctacgcgcgcaacacatcctactgtggggccgctcacg 606  
QY 301 ctgaactgtcccccgcgctctgcttcgacgggagggagggcttcttcggcactcaact 360  
Db 607 ctgaactgtcccccgcgctctgcttcgacgggagggagggcttcttcggcactcaact 666  
QY 361 gcgtccgtgctgagcctcctcggcctacgcgctgagcgagcagcctcacatggcgcgag 420  
Db 667 gcgtccgtgctgagcctcctcggcctacgcgctgagcgagcagcctcacatggcgcgag 726  
QY 421 gggcccgcccgctcctcagtcggggcgacgctgagcgatggcagcggcggcctggggc 480  
Db 727 gggcccgcccgctcctcagtcggggcgacgctgagcgatggcagcggcggcctggggc 786  
QY 481 gtgtcgtgctcctcggcctcctcgcagcgtggcctgggaattgcctgggtgcctgggac 540  
Db 787 gtgtcgtgctcctcggcctcctcgcagcgtggcctgggaattgcctgggtgcctgggac 846  
QY 541 gcttgctccactgtcttggcggctctctacgccaagcctacgtgctcttctgcgtgcgcgc 600  
Db 847 gcttgctccactgtcttggcggctctctacgccaagcctacgtgctcttctgcgtgcgcgc 906













KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;  
XX neuroprotective; ds.  
OS Homo sapiens.  
XX WO200136473-A2.  
XX 25-MAY-2001.  
XX 16-NOV-2000; 2000WO-US31581.  
XX 16-NOV-1999; 99US-0165838.  
XX 17-NOV-1999; 99US-0166071.  
XX 19-NOV-1999; 99US-0166678.  
XX 28-DEC-1999; 99US-0173396.  
XX 22-FEB-2000; 2000US-0184129.  
XX 28-FEB-2000; 2000US-0185421.  
XX 28-FEB-2000; 2000US-0185554.  
XX 03-MAR-2000; 2000US-0186530.  
XX 09-MAR-2000; 2000US-0186811.  
XX 17-MAR-2000; 2000US-0190310.  
XX 21-MAR-2000; 2000US-0190800.  
XX 20-APR-2000; 2000US-0198568.  
XX 02-MAY-2000; 2000US-0201190.  
XX 08-MAY-2000; 2000US-0203111.  
XX 25-MAY-2000; 2000US-0207094.  
XX (PHAA ) PHARMACIA & UPJOHN CO.  
XX Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J;  
XX Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;  
XX WPI; 2001-389826/41.  
XX P-PSDB; AAG80933.  
XX New G protein-coupled receptor (nGPCR-x) and its encoding  
XX polynucleotide useful for diagnosing and treating e.g. schizophrenia -  
XX Claim 4; Page 77; 261pp; English.  
XX The present invention relates to novel G protein-coupled receptors  
XX (nGPCRx; where x is 1, 2, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,  
XX 28, 31-38, 40, 41, 53-60) and their coding sequences. The present  
XX sequence is the coding sequence for one such G protein-coupled receptor.  
XX GPCRs are also known as seven transmembrane receptors and function in  
XX signal transduction. The nGPCRx coding sequences are useful for  
XX screening a human to diagnose a disorder affecting the brain or a genetic  
XX predisposition, specifically schizophrenia. nGPCRx are useful for  
XX identifying compounds useful for treating schizophrenia. Detection of  
XX nGPCRx in a sample is useful as a diagnostic tool for diseases or  
XX disorders e.g. thyroid disorders, renal failure, rheumatoid arthritis,  
XX CNS disorders, infections such as HIV-1, metabolic and cardiovascular  
XX diseases, proliferative disorders and hormonal disorders. Modulators of  
XX nGPCRx activity have the utility for treating neurological disorders,  
XX including schizophrenia, ADHD/ADD (attention deficit-hyperactivity  
XX disorder/attention deficit disorder), and neuronal disorders such as  
XX Alzheimer's disease, Parkinson's disease, migraine and senile dementia.  
XX Additional disorders include inflammatory conditions (e.g. Crohn's  
XX disease), rheumatoid arthritis, autoimmune disorders, cancers,  
XX respiratory ailments such as asthma, and inflammatory diseases e.g.  
XX inflammatory bowel disease.  
XX Sequence 270 BP; 30 A; 101 C; 92 G; 47 T; 0 other;  
Query Match. 22.4%; Score 268.4; DB 22; Length 270;  
Best Local Similarity 99.6%; Pred. No. 4.1e-42;  
Matches 269; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 208 cccatgttctctctgagcagctcagctgtgtcgtatctgtgagcgcgcgcctac 267  
DB 1 cccatgttctctctgagcagctcagctgtgtcgtatctgtgagcgcgcgcctac 60

QY 268 gccccaacatctactgtcggggcgctcacgtgaaactgtccccgcgctctgttc 327  
DB 61 gccccaacatctactgtcggggcgctcacgtgaaactgtccccgcgctctgttc 120  
QY 328 gcacggaggagggcgcttctcgttgacactactcgtcgtcgtgagcctctcgttc 387  
DB 121 gcacggaggagggcgcttctcgttgacactactcgtcgtcgtgagcctctcgttc 180  
QY 388 gcgctggagcgcagcctcaccatggcgagggggcgccgcgctctccagtcgggg 447  
DB 181 gcgctggagcgcagcctcaccatggcgagggggcgccgcgctctccagtcgggg 240  
QY 448 gcgctggagcgcagcctcaccatggcgagggggcgccgcgctctccagtcgggg 477  
DB 241 gcgctggagcgcagcctcaccatggcgagggggcgccgcgctctccagtcgggg 270  
RESULT 12  
AAT58506  
ID AAT58506 standard; DNA; 2232 BP.  
XX  
AC AAT58506;  
XX  
DT 01-APR-1997 (first entry).  
XX  
DE cDNA encoding p(rat-edg), G-protein coupled receptor.  
XX  
KW p(H218); G-protein coupled receptor; cell differentiation; proliferation;  
KW proline directed kinase; cell division; growth factor response; rat-edg;  
KW therapy; diagnosis; ss.  
XX  
OS Rattus rattus.  
XX  
FH Key Location/Qualifiers  
FT CDS 269..1420  
FT misc\_signal /tag= a  
FT 1768..1772  
FT /tag= b  
FT /note= "associated with mRNA degradation"  
XX  
XX US5585476-A.  
PN 17-DEC-1996.  
PD 15-FEB-1994; 94US-0196989.  
PF 15-FEB-1994; 94US-0196989.  
PR 15-FEB-1994; 94US-0196989.  
XX (MACL/) MACLENNAN A J.  
XX MacLennan AJ;  
XX WPI: 1997-051235/05.  
XX P-PSDB; AAW01664.  
XX  
XX DNA encoding rat protein p(H218) - associated with cell  
XX proliferation and/or differentiation  
XX Example 9; Column 19-24; 33pp; English.  
XX  
XX This sequence is a novel rat cDNA which encodes p(rat-edg), a member  
XX of the G-protein coupled receptor superfamily of proteins. The amino  
XX acid similarity between p(H218) (AAW01663) and p(rat-edg) (AAW01664)  
XX suggests that they may be activated by the same endogenous ligand(s). The  
XX expression pattern of mRNA transcripts of both genes in cell lines,  
XX various rat tissues and developing rat brain suggests that they both play  
XX a role in cell proliferation and/or differentiation.  
SQ Sequence 2232 BP; 552 A; 614 C; 488 G; 578 T; 0 other;  
Query Match 21.9%; Score 261.6; DB 18; Length 2232;











GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein : protein search, using sw model

Run on: January 16, 2002, 22:38:38 ; Search time 73.84 Seconds  
(without alignments)  
399.257 Million cell updates/sec

Title: US-09-842-316-2

Perfect score: 2019

Sequence: 1 MESGLLRAPVSEVIVLHYN.....TGSPGAPTAARTLVSPAAD 398

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_1101.\*  
1: /SIDS2/gcgdata/geneseq/geneseq/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2019	100.0	398	22	Human nGPCR9 #2.
2	2019	100.0	398	22	Amino acid sequenc
3	2019	100.0	399	22	Human PFI-006 prot
4	2018	100.0	398	21	Amino acid sequenc
5	2013	99.7	500	21	Human orphan G pro
6	2013	99.7	500	21	Human G protein co
7	1714.5	84.9	400	22	Rat G protein-coup
8	865.5	42.9	382	20	Human EDG-2 protel
9	865.5	42.9	382	22	LPA receptor-relat
10	863.5	42.8	382	21	Mouse EDG1 polypep
11	853.5	42.3	383	18	p(rat-edg), G-prot

12	853.5	42.3	383	20	AAW87791	Rat-edg, G-protein
13	842	41.7	381	20	AAAY3205	Human EDG-1c recep
14	834	41.3	381	12	AAAI4337	G-protein coupled
15	834	41.3	381	22	AAAB70503	Angiogenesis prote
16	776	38.4	334	15	AAAR48752	Human endothelial
17	776	38.4	334	17	AAAW02724	Human endothelial
18	774.5	38.4	378	20	AAAY05490	Human EDG-3 protel
19	774.5	38.4	378	20	AAAY59176	Human EDG3sb polyp
20	774.5	38.4	378	22	AAAU00303	LPA receptor-relat
21	750.5	37.2	378	22	AAAB71751	Murine EDG3 polype
22	733.5	36.3	379	22	AAAB73483	Rat G protein-coup
23	732.5	36.3	352	15	AAAR58712	Fragment of recept
24	732.5	36.3	352	18	AAAW01663	p(H218), G-protein
25	732.5	36.3	352	20	AAAY05492	Human EDG-5 protel
26	732.5	36.3	352	20	AAW87790	Human EDG-5 protel
27	732.5	36.3	352	22	AAAU00305	LPA receptor-relat
28	728.5	36.1	353	20	AAAY49904	Rat H218, G-protein
29	722	35.8	353	21	AAAB18889	Human EDG family H
30	721.5	35.7	353	20	AAAY28288	Amino acid sequenc
31	721.5	35.7	353	20	AAAY28289	Predicted polypept
32	720	35.7	353	20	AAAY49905	EDG-4 amino acid s
33	712.5	35.3	352	21	AAAB03966	Human EDG family H
34	647	32.0	384	20	AAAY06412	Murine EDG5 polype
35	645.5	32.0	384	20	AAAY06411	Human EDG-7 recept
36	645.5	32.0	384	21	AAAY69500	A human G-protein
37	645.5	32.0	384	21	AAW90862	Human edg6 protein
38	607.5	30.1	386	21	AAW90863	Murine edg6 protel
39	548	27.1	509	19	AAW80955	Amino acid sequenc
40	518.5	25.7	364	19	AAW85451	Rat olfactory bulb
41	518.5	25.7	364	19	AAW59501	Rat olfactory bulb
42	518.5	25.7	364	20	AAW80990	Murine lysophospha
43	518.5	25.7	364	22	AAAB7347	Murine LPA recept
44	508.5	25.2	351	20	AAW82659	Human edg-6 protel
45	507.5	25.1	364	18	AAW07619	Human G-protein co

#### ALIGNMENTS

RESULT 1  
AAG80967  
ID AAG80967 standard; Protein; 398 AA.

XX AAG80967;

XX 28-AUG-2001 (first entry)

DT Human nGPCR9 #2.

DE G protein-coupled receptor; nGPCR; seven transmembrane receptor;  
KW signal transduction; schizophrenia; thyroid disorder; renal failure;  
KW rheumatoid arthritis; CNS disorder; infection; metabolic disease;  
KW cardiovascular disease; proliferative disorder; hormonal disorder;  
KW neurological disorder; neuronal disorder; Alzheimer's disease;  
KW attention deficit-hyperactivity disorder/attention deficit disorder;  
KW Parkinson's disease; migraine; senile dementia; inflammatory disease;  
KW rheumatoid arthritis; autoimmune disorder; respiratory ailment;  
KW neuroprotective.

XX Homo sapiens.

XX WO200136473-A2.

XX PD 25-MAY-2001.

XX 16-NOV-2000; 2000WO-US31581.

XX 16-NOV-1999; 99US-0165838.

XX 17-NOV-1999; 99US-0166071.

XX 19-NOV-1999; 99US-0166678.

XX 28-DEC-1999; 99US-0173396.

XX 22-FEB-2000; 2000US-0184129.

XX 28-FEB-2000; 2000US-0185421.

28-FEB-2000; 2000US-0185554.  
02-MAR-2000; 2000US-0186530.  
03-MAR-2000; 2000US-0186811.  
09-MAR-2000; 2000US-0188114.  
17-MAR-2000; 2000US-0190310.  
21-MAR-2000; 2000US-0190800.  
20-APR-2000; 2000US-0198568.  
02-MAY-2000; 2000US-0201190.  
08-MAY-2000; 2000US-0203111.  
23-MAY-2000; 2000US-0207094.  
(PHAA ) PHARMACIA & UPJOHN CO.  
Vogeli G, Wood LS, Parodi LA, Hiebsch RR, Lind P, Slightom J,  
Schellin KA, Kaytes PS, Bannigan CM, Ruff V, Sejlitz T, Huff RM;  
WPI; 2001-389826/41.  
N-PSDB; AAH51007.  
New G protein-coupled receptor (nGPCR-x) and its encoding  
polynucleotide useful for diagnosing and treating e.g. schizophrenia -  
Claim 37; Page 89; 261pp; English.  
The present invention relates to novel G protein-coupled receptors  
(nGPCRx; where x is 1, 3, 4, 5, 9, 11, 12, 14-18, 20, 21, 22, 24, 27,  
28, 31-38, 40, 41, 53-60) and their coding sequences. The present  
sequence is one such G protein-coupled receptor. GPCRs are also known as  
seven transmembrane receptors and function in signal transduction. The  
nGPCRx coding sequences are useful for screening a human to diagnose a  
disorder affecting the brain or a genetic predisposition, specifically  
schizophrenia. nGPCRx are useful for identifying compounds useful for  
treating schizophrenia. Detection of nGPCRx in a sample is useful as a  
diagnostic tool for diseases or disorders e.g. thyroid disorders, renal  
failure, rheumatoid arthritis, CNS disorders, infectious such as HIV-1,  
metabolic and cardiovascular diseases, proliferative disorders and  
hormonal disorders. Modulators of nGPCRx activity have the utility for  
treating neurological disorders, including schizophrenia, ADHD/ADD  
(attention deficit-hyperactivity disorder/attention deficit disorder),  
and neuronal disorders such as Alzheimer's disease, Parkinson's disease,  
migraine and senile dementia. Additional disorders include inflammatory  
conditions (e.g. Crohn's disease), rheumatoid arthritis, autoimmune  
disorders, cancers, respiratory ailments such as asthma, and inflammatory  
diseases e.g. inflammatory bowel disease.  
Sequence 398 AA;  
Query Match 100.0%; Score 2019; DB 22; Length 398;  
Best Local Similarity 100.0%; Pred. No. 8.5e-181;  
Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MESGLRPAVSEVIVLHNYTKRGARYQFGAGLRADAVVCLAVCAFIENLAVLLV 60  
DB 1 mesglrpaavsevivlhyntgkrgarypgaglradaavvcclavcafienvlavl 60  
QY 61 LGRHPRFAPMELLGSLTSLDLAGAAYANILLGSLTLKLSPALWFAREGGVFALT 120  
DB 61 lgrhprfapmfillgsitclsdllagaayaanillsgptlklspalwfareggvvalt 120  
QY 121 ASVLSLALERSLTARRGPAPVSSRGRTLAMAAAWGVSLLGLLPALGNCLGRLD 180  
DB 121 asvlsllalersltarrgipapvssrgrtlamaaawgvsllgllpalgnclgrld 180  
QY 181 ACSTVPLVAKAVLCVLAFCVLAATCALYARIYQVQANARRLPARPGTAGTSTRA 240  
DB 181 acstvlplyakavlfvcvllaatcalaryiqvqanarrlpargtagtsttra 240  
QY 241 RRKPRSLALRLTSVVLAFVACMGPFLLLLLDVACPARTCPVLLQADPFLGLMANSL 300  
DB 241 rrpksrslalrltsvvlafvacwgpflillldvacpartcpvllqadpflglman 300  
QY 301 LNPFIYTLNRDLRHALLRLVCCGRHSCGRDPGSGQSASAAEASGGLRCLPPGLDGSF 360

DB 301 lnpfiytlnrdlrhailrlvccgrhscgrdpsgqsasaaesgglrrclppgidgsf 360  
QY 361 SSSRSRSPQRGLDTSGTSGPGAPTAARTLVSEPAAD 398  
DB 361 sgsrsrsprqrgldtsgtsgpgaptaartlvsepaad 398  
RESULT 2  
AAB31665  
ID AAB31655 standard; Protein: 398 AA.  
XX  
AC AAB31655;  
XX  
DT 30-APR-2001 (first entry)  
XX  
DE Amino acid sequence of a human AXOR29 polypeptide.  
XX  
KW Human; G-coupled protein receptor; AXOR29; infection; HIV-1; HIV-2; pain;  
KW cancer; diabetes; obesity; anorexia; bulimia; asthma; hypotension;  
KW parkinson's disease; acute heart failure; hypertension; angina pectoris;  
KW urinary retention; osteoporosis; myocardial infarction; stroke; ulcer;  
KW allergy; benign prostatic hypertrophy; migraine; vomiting; schizophrenia;  
KW psychotic disorder; neurological disorder; anxiety; manic depression;  
KW depression; delirium; dementia; mental retardation; vaccine.  
XX  
OS Homo sapiens.  
XX  
PN MO200104119-A2  
XX  
PD 18-JAN-2001  
XX  
PF 13-JUL-2000; 2000WO-US19001.  
XX  
PR 13-JUL-1999; 99GB-0016417.  
PR 08-DEC-1999; 99US-0169573.  
XX  
PA (SMIK ) SMITHKLINE BEECHAM CORP.  
PA (SMIK ) SMITHKLINE BEECHAM PLC.  
PI Ames RS, Elshourbagy N, Foley JJ, Michalovich D, Sarau HM;  
PI Smith R, Tsui P, Vawter L, Agarwal P, Lane P;  
XX WPI; 2001-147176/15.  
DR N-PSDB; AAF25128.  
XX  
PT Novel G-coupled protein receptor, AXOR29 useful for treating diseases  
PT such as microbial infections, cancers, obesity, asthma, diabetes,  
PT hypotension, osteoporosis, myocardial infarction and neurological  
PT disorders  
XX  
PS Claim 1; Fig 2; 41pp; English.  
XX  
CC The present sequence represents a human G-coupled protein receptor,  
CC designated AXOR29. AXOR29 polypeptides and polynucleotides are useful  
CC for treating and diagnosing infections such as bacterial, fungal,  
CC protozoan and viral infections, particularly infections caused by  
CC HIV-1 or HIV-2, pain, cancers, diabetes, obesity, anorexia, bulimia,  
CC asthma, parkinson's diseases, acute heart failure, hypotension,  
CC hypertension, urinary retention, osteoporosis, angina pectoris,  
CC myocardial infarction, stroke, ulcers, allergies, benign prostatic  
CC hyper trophy, migraine, vomiting, psychotic and neurological disorders  
CC including anxiety, schizophrenia, manic depression, depression, delirium,  
CC dementia and severe mental retardation. the polypeptide is also useful  
CC as a vaccine for inducing immunological response in a mammal.  
XX  
SQ Sequence 398 AA;

Query Match 100.0%; Score 2019; DB 22; Length 398;  
Best Local Similarity 100.0%; Pred. No. 8.5e-181;  
Matches 398; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESGLLRPAVSEVIVLHNYTGLRGARYQPGAGLRADAVVCLAVCAFIYLENLAVLLV 60  
Db 1 mesgllrpapvsevivlhyntgkrgaryqpgaglradvavclavcafiylenlavllv 60  
QY 61 LGRHPRHAPMFLLSITLSDLLAGAAAYANILLSGPLTKLSPALWFAREGGVFALT 120  
Db 61 lgrhprhaphmflfllsgtldslagaaayaanillsgpltklspalwfaregvgfvait 120  
QY 121 ASVLSLLAIALERSLTMRGAPVSSRGRTLMAAAAGVSLLLGLLPAWGNCILGRD 180  
Db 121 asvlsllaialersltmarrgppvssrgtrtlamaaaagvslillglpalgwncilgrld 180  
QY 181 ACSIVPLKAYAVLFCVLAFAVGLAALICARYICQVRANRRPARGTAGTSTRA 240  
Db 181 acstivplkayavlfvcvlfavglfaalicaaryicqvrannrrparpgtagttstra 240  
QY 241 RRPRLSALLRTLNVLLAFVACWGPLFLLLLDVPACPARTCPVLLOADPFLGLMANSL 300  
Db 241 rrprrsallrtlnvllafvacwgpflflllldvpacpartcpvllqadpflglmansl 300  
QY 301 LNPITVLTNRDLRHALLRVCCGRHSCGRDPSSGQSSAASAGGLRCLPPGLDGSF 360  
Db 301 lnpitvltlnrdlrhallrvccgrhscgrdpssgqssasaasagglrcclppgldgsf 360  
QY 361 SRSERSPQRDGLDTSSTGSGPAPTARTLVSEPAAD 398  
Db 361 srserspqrldtsgtsgspgaptartlvsepaad 398  
RESULT 3  
ID AAB74956 standard; Protein; 399 AA.  
XX AAB74956;  
DT 06-JUL-2001 (first entry)  
DE Human PFI-006 protein sequence SEQ ID NO:2.  
XX Human; PFI-006: G-protein coupled receptor; GPCR; chromosome 19;  
KW anorectic; antidiabetic; neuroprotective; neuroleptic; nootropic;  
KW psychotropic; antiinflammatory; cytostatic; vulnerary; dermatological;  
KW osteopathic; cardiovascular; antiallergic; gene therapy; diabetes;  
KW G-protein coupled receptor modulator; signal transduction; obesity;  
KW metabolic disease; neurological disease; psychotherapeutic; inflammation;  
KW urogenital disease; reproduction; sexual medicine; cancer; tissue repair;  
KW dermatology; skin pigmentation; photoaging; frailty; osteoporosis;  
KW cardiovascular disease; gastrointestinal disease; antinfection; allergy;  
KW respiratory disease; sensory organ disorder; sleep disorder; hair loss.  
OS Homo sapiens.  
XX  
XX EPI090925-AL.  
PD 11-APR-2001.  
PF 06-OCT-2000; 2000EP-0308853.  
PR 08-OCT-1999; 99GB-0023890.  
PA (PFI2 ) PFIZER LTD.  
PA (PFI2 ) PFIZER INC.  
PI Harland L;  
XX  
DR WPI; 2001-302043/32.  
DR N-PSDB; AAF87722.  
XX  
PT New human G-protein coupled receptor polynucleotide and polypeptide,  
PT useful in drug screening, as well as for diagnosing or treating  
PT diseases associated with signal transduction, e.g. obesity,  
PT inflammation, tissue repair, or cancer

PS  
XX  
CC The present sequence is human PFI-006 which is a putative G-protein  
CC coupled receptor (GPCR). PFI-006 was identified from chromosome 19.  
CC PFI-006 has anorectic, antidiabetic, neuroprotective, neuroleptic,  
CC nootropic, psychotropic, antiinflammatory, cytostatic, vulnerary,  
CC dermatological, osteopathic, cardiovascular and antiallergic activities.  
CC It can be used in pharmaceuticals or in the manufacture of medicaments.  
CC for the treatment of a patient having need to modulate, i.e. agonise or  
CC (selectively) antagonise PFI-006 protein. A compound which modulates  
CC PFI-006 can be used in gene therapy. PFI-006 polynucleotides and  
CC proteins can be used in diagnosing and treating a variety of diseases.  
CC These are also useful for evaluating and/or screening agents that can  
CC modulate the GPCR. In particular, they are useful for screening drug  
CC candidates for treating diseases associated with signal transduction,  
CC including obesity, diabetes and metabolic disease, neurological disease,  
CC psychotherapeutics, urogenital disease, reproduction and sexual  
CC medicine, inflammation, cancer, tissue repair, dermatology, skin  
CC pigmentation, photoaging, frailty, osteoporosis, cardiovascular disease,  
CC gastrointestinal disease, antinfection, allergy and respiratory  
CC disease, sensory organ disorders, sleep disorders or hair loss.  
XX  
SQ Sequence 399 AA;

Query Match 100.0%; Score 2019; DB 22; Length 399;  
Best Local Similarity 100.0%; Pred. No. 8.5e-181;  
Matches 398; Conservative. 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MESGLLRPAVSEVIVLHNYTGLRGARYQPGAGLRADAVVCLAVCAFIYLENLAVLLV 60  
Db 1 mesgllrpapvsevivlhyntgkrgaryqpgaglradvavclavcafiylenlavllv 60  
QY 61 LGRHPRHAPMFLLSITLSDLLAGAAAYANILLSGPLTKLSPALWFAREGGVFALT 120  
Db 61 lgrhprhaphmflfllsgtldslagaaayaanillsgpltklspalwfaregvgfvait 120  
QY 121 ASVLSLLAIALERSLTMRGAPVSSRGRTLMAAAAGVSLLLGLLPAWGNCILGRD 180  
Db 121 asvlsllaialersltmarrgppvssrgtrtlamaaaagvslillglpalgwncilgrld 180  
QY 181 ACSIVPLKAYAVLFCVLAFAVGLAALICARYICQVRANRRPARGTAGTSTRA 240  
Db 181 acstivplkayavlfvcvlfavglfaalicaaryicqvrannrrparpgtagttstra 240  
QY 241 RRPRLSALLRTLNVLLAFVACWGPLFLLLLDVPACPARTCPVLLOADPFLGLMANSL 300  
Db 241 rrprrsallrtlnvllafvacwgpflflllldvpacpartcpvllqadpflglmansl 300  
QY 301 LNPITVLTNRDLRHALLRVCCGRHSCGRDPSSGQSSAASAGGLRCLPPGLDGSF 360  
Db 301 lnpitvltlnrdlrhallrvccgrhscgrdpssgqssasaasagglrcclppgldgsf 360  
QY 361 SRSERSPQRDGLDTSSTGSGPAPTARTLVSEPAAD 398  
Db 361 srserspqrldtsgtsgspgaptartlvsepaad 398

RESULT 4  
AAY69384  
ID AAY69384 standard; Protein; 398 AA.  
XX  
AC AAY69384;  
XX  
DT 19-JUN-2000 (first entry)  
XX  
DE Amino acid sequence of a 14274 receptor protein.  
XX  
KW G-protein coupled receptor; EDG receptor; 14274 receptor;  
KW spleen disorder; splenomegaly; lung disorder; emphysema;  
KW chronic bronchitis; colon disorder; atresia; stenosis; enterocolitis;  
KW liver disorder; jaundice; cholestasis; brain disorder; cerebral edema;  
KW cerebrovascular disease; Alzheimer's disease; Parkinson's disease;

KW T-cell disorder; autoimmune disease; transplant rejection; skin disease;  
 KW vitiligo; lentigo; malignant melanoma; heart disorder; breast disorder;  
 KW ischemic heart disease; kidney disorder; cystic disease; inflammation;  
 KW mastitis; mammary duct ectasia; prostate disorder; benign enlargement;  
 KW tumour.

OS Homo sapiens.

XX Key Location/Qualifiers

PH 1..39 "extracellular domain"  
 FT 40..62  
 FT /note- "transmembrane domain segment"  
 FT 71..95  
 FT /note- "transmembrane domain segment"  
 FT 114..131  
 FT /note- "transmembrane domain segment"  
 FT 132..134  
 FT /note- "GPCR signature"  
 FT 152..173  
 FT /note- "transmembrane domain segment"  
 FT 192..213  
 FT /note- "transmembrane domain segment"  
 FT 253..273  
 FT /note- "transmembrane domain segment"  
 FT 291..308  
 FT /note- "transmembrane domain segment"  
 FT 309..398  
 FT /note- "intracellular domain"

PN WO200011166-A1.

XX 02-MAR-2000.

XX 19-AUG-1999; 99WO-US18976.

XX 19-AUG-1998; 98US-0136726.

XX 19-AUG-1999; 99US-0377429.

XX (MILL-) MILLENNIUM PHARM INC.

XX Glucksmann MA, Welch NS, Hunter JJ;

XX WPI: 2000-224699/19.

XX N-PSDB: AA261489.

XX New G-protein coupled receptor used in receptor assays as a target for  
 diagnosis and treatment of receptor-mediated disorders of the spleen,  
 lung, brain, heart and liver.

PS Claim 1; Fig 1; 102pp; English.

CC The present sequence represents a polypeptide which is a G-protein  
 CC coupled receptor (GPCR) related to the EDG (undefined) receptor family,  
 CC and is designated the 14274 receptor. The receptor polypeptides are  
 CC useful for producing antibodies specific for the GPCR 14274 receptor  
 CC protein, and in drug screening assays, in cell-based or cell-free  
 CC systems. The receptor polypeptide may also be used to identify compounds  
 CC that modulate receptor activity, that is compounds which stimulate or  
 CC inhibit interaction between the receptor protein and a target molecule.  
 CC Modulators of receptor protein activity are used to treat subjects with  
 CC a disorder mediated by the GPCR 14274 receptor pathway, such as;  
 CC disorders of the spleen (such as splenomegaly), lung disorders (such  
 CC as emphysema, chronic bronchitis), disorders of the colon (such as  
 CC atresia, stenosis, enterocolitis), disorders of the liver (such as  
 CC jaundice, cholestasis), disorders of the brain (such as cerebral edema,  
 CC cerebrovascular disease, Alzheimer's disease, Parkinson's disease),  
 CC disorders involving T-cells (such as autoimmune diseases, transplant  
 CC rejection), skin diseases (such as vitiligo, lentigo, malignant  
 CC melanoma), heart disorders (such as ischemic heart disease), kidney  
 CC disorders (such as cystic diseases), disorders of the breast (such as  
 CC inflammation, mastitis, mammary duct ectasia), and prostate disorders  
 CC (such as inflammations, benign enlargement, tumours).

XX	Sequence	398 AA;
SQ	Query Match	100.0%; Score 2018; DB 21; Length 398;
	Best Local Similarity	99.7%; Pred. No. 1..1e-180;
	Matches 397; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	

  

QY	1	NESGLLRPAVSEVIVLHYNTYKLRGARYQFAGLRADAVVCLAVCAFIIVLENLAVLV	60
DB	1	mesgllrpapvsevivlhyntgkrgaryqpgaglradvclavcafiivlenlavlv	60
QY	61	LGHRPFRHAPMELLSLTSLDLAGAAYANILLSGPLTLKSLPALWFREGGVFVALT	120
DB	61	lgrhprfhapmllsgltisdllagaayaanillsgpltlklsplwfareggvfvalt	120
QY	121	ASVLSLAIALERSLTMARPGAPVSSRGRTLMAAAAGVSLIIGLLPALGWNCGLRLD	180
DB	121	asvlsilalalersltmarpgapvssrgrtlamaaagvslilgllpalgwnclgrld	180
QY	181	ACSTVLPYAKAYVFCVLAFAVGIILAAICALYARIYCOVRANARRLPARPCTAGTTSTRA	240
DB	181	acstvlpayakayvfcvlfafvgilaicalyariycqiranarrlparpctagttstra	240
QY	241	RRKPSLALLRTLSSVLLAFVACWGPLFLLLLLDVACPARTCPVLQADPFLGLAMANS	300
DB	241	rrkprslallrtlsvllafvacwgpflfllldvacpartcpvlqadpflglamansl	300
QY	301	LNPIYITLNRDLRHALLRLVCCGRHSCGRDPGSGQSAASAAESGGLRRLCPPLDGSF	360
DB	301	lhpilyltlnrdlrhallrlvccgrhscgrdpsgsgqsasaaesgglrrclppldgsf	360
QY	361	SGSERSSQRCGLDTSSTGSPGAPTAAARTLVSEPAAD	398
DB	361	sgserssqrcldtsgtsgpaptartlvsepaad	398

  

RESULT	5
AAV71305	
ID	AAV71305 standard; Protein; 500 AA.
XX	AAV71305;
AC	AAV71305;
XX	
DT	02-NOV-2000 (first entry)
XX	
DE	Human orphan G protein-coupled receptor hCHN6.
XX	
KW	Human; orphan G protein-coupled receptor; GPCR; hCHN6; drug screening;
KW	transmembrane receptor; expressed sequence tag; EST; signal cascade.
XX	
OS	Homo sapiens.
XX	
PN	WO20001258-A2.
XX	
PD	02-JUN-2000.
XX	
PF	13-OCT-1999; 99WO-US23687.
XX	
PR	20-NOV-1998; 98US-0109213.
PR	16-FEB-1999; 99US-0120416.
PR	26-FEB-1999; 99US-0121852.
PR	12-MAR-1999; 99US-0123946.
PR	12-MAR-1999; 99US-0123949.
PR	28-MAY-1999; 99US-0136436.
PR	28-MAY-1999; 99US-0136437.
PR	28-MAY-1999; 99US-0136439.
PR	28-MAY-1999; 99US-0136567.
PR	28-MAY-1999; 99US-0137127.
PR	28-MAY-1999; 99US-0137131.
PR	29-JUN-1999; 99US-0141448.
PR	29-SEP-1999; 99US-0156555.
PR	29-SEP-1999; 99US-0156633.
PR	29-SEP-1999; 99US-0156634.

29-SEP-1999; 99US-0156653.  
 01-OCT-1999; 99US-0157280.  
 01-OCT-1999; 99US-0157281.  
 01-OCT-1999; 99US-0157282.  
 01-OCT-1999; 99US-0157293.  
 01-OCT-1999; 99US-0157294.  
 01-OCT-1999; 99US-0157294.  
 12-OCT-1999; 99US-0416760.  
 12-OCT-1999; 99US-0417044.  
 (AREN-) ARENA PHARM INC.  
 Chen R, Dang HT, Llaw CW, Lin I;  
 N-PSDB; AAD01132.  
 Novel human orphan G protein-coupled receptors and the encoding cDNAs  
 for use in the identification of G protein-coupled receptor agonists -  
 Claim 58; Page 79-81; 102pp; English.  
 The present amino acid sequence is the hCHN6, an endogenous human  
 orphan G protein-coupled receptor (GPCR). The hCHN6 cDNA was identified  
 using ESTs (expressed sequence tag) AA804531 and 2134670 as a probe.  
 The orphan GPCR of the invention, like all GPCRs has seven transmembrane  
 alpha helices with an extracellular N-terminus and an intracellular  
 C-terminus. However, no endogenous ligands has yet been identified for  
 the proteins of the invention. The orphan GPCRs may be used in the  
 identification of their endogenous ligands, and to screen potential GPCR  
 agonists and antagonists for use as pharmaceutical agents. The proteins  
 may also be used in the study of GPCR-mediated signalling cascades, and  
 to elucidate their precise role in normal and diseased human conditions.  
 CC Nucleic acid encoding human orphan GPCRs may be used for tissue  
 CC localisation expression analysis to provide information about their  
 CC function in healthy and pathological states.  
 XX Sequence 500 AA;  
 SQ  
 Query Match 99.7%; Score 2013; DB 21; Length 500;  
 Best Local Similarity 99.7%; Pred. No. 4.1e-180;  
 Matches 397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MESGLRPAVSEVIVLHNYTGKIRGARYQPCGRLRADAVVCLAVCAFIENLAVLLV 60  
 Db 103 mesgllrpapvsevivlhyntgkrgasyqpgaglradaavclavcafienvlavllv 162  
 QY 61 LGRHPRFAPMFLLIGSLTSLDLAGAAYAAANILLGPTLKLSPALWFAREGGVFAVLT 120  
 Db 163 lgrhprfhapmflligsltdlagaayaanillsgptlklspalwfareggvfvavlt 222  
 QY 121 ASVLSLATALERSLTMRGAPVSSRGRTLMAAAANGVSLLLGLLPALGNCLGRLD 180  
 Db 223 asvlsllatalersltmrargpavssrgtrtlamaaaangvslllgllpalgncldgrld 282  
 QY 181 ACSTVPLKAYKAVLCVLAFCVGLAALCALYARYICVVRANRRLPARPCAGTCTSTRA 240  
 Db 283 acstvlplykayvlfcvlaflavglalaaicalaryicvvrannrripapcagcttstra 342  
 QY 241 RRPKRLALRLTSLVLLAFACVAGPLFLLLLLDVACPARTCPVLLQADPFLGAMANS 300  
 Db 343 rrpkrallrltslvllafavcwgplfllllldvacpartcpvllqadpflgiamansl 402  
 QY 301 LNPPIITLNRDLRALRLVCCGRHSGRPSGSGQSOASAAEASGGLRCLLPGLDGSF 360  
 Db 403 lnpilvllcnrdlrallrvccgrhscgrpsgsgsqsoasaaeasgglrrcplpdlgsf 462  
 QY 361 SGSESSPQDGLDTSVGTSGSPGAPTAARTLVSEPAAD 398  
 Db 463 sgseerspqrldtsgtsgspgaptartlvsepaad 500

RESULT 6

AA802839  
 ID AAB02839 standard; Protein; 500 AA.  
 XX  
 AC AAB02839;  
 XX  
 DT 22-AUG-2000 (first entry)  
 XX  
 DE Human G protein coupled receptor hCHN6 protein SEQ ID NO:32.  
 XX  
 KW Human; G protein coupled receptor; GPCR; transmembrane receptor;  
 KW identification; agonist; screening; therapeutic; pharmaceutical;  
 KW mutant.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2000022131-A2  
 XX  
 PD 20-APR-2000.  
 XX  
 PF 13-OCT-1999; 99WO-US24065.  
 XX  
 PR 13-OCT-1998; 98US-0170496.  
 PR 12-NOV-1998; 98US-0108029.  
 PR 27-NOV-1998; 98US-0109213.  
 PR 16-FEB-1999; 98US-0110060.  
 PR 26-FEB-1999; 98US-0120416.  
 PR 12-MAR-1999; 98US-0121852.  
 PR 12-MAR-1999; 98US-0123944.  
 PR 12-MAR-1999; 98US-0123945.  
 PR 12-MAR-1999; 98US-0123946.  
 PR 12-MAR-1999; 98US-0123948.  
 PR 12-MAR-1999; 98US-0123949.  
 PR 12-MAR-1999; 98US-0123951.  
 PR 28-MAY-1999; 98US-0136436.  
 PR 28-MAY-1999; 98US-0136437.  
 PR 28-MAY-1999; 98US-0136439.  
 PR 28-MAY-1999; 98US-0137127.  
 PR 28-MAY-1999; 98US-0137131.  
 PR 28-MAY-1999; 98US-0137567.  
 PR 30-JUN-1999; 98US-0141448.  
 PR 27-AUG-1999; 98US-0151114.  
 PR 03-SEP-1999; 98US-0152524.  
 PR 29-SEP-1999; 98US-0156633.  
 PR 29-SEP-1999; 98US-0156555.  
 PR 29-SEP-1999; 98US-0156634.  
 XX  
 (AREN-) ARENA PHARM INC.  
 Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;  
 PI Gore M, Llaw CW, Lin I, Lowitz K, White C;  
 XX  
 WPI: 2000-317986/27.  
 DR N-PSDB; AAA46033.  
 XX  
 PT Non-endogenous, human G protein-coupled receptors for screening  
 PT receptor, inverse or partial agonists useful as therapeutic agents  
 XX  
 PS Example 1; Page 109-111; 187pp; English.  
 XX  
 CC The present invention describes transmembrane receptors, preferably  
 CC human G protein coupled receptors (GPCR), for which the endogenous  
 CC ligand is unknown (orphan GPCR receptors). More specifically the present  
 CC invention relates to non-endogenous, constitutively activated versions  
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for  
 CC the direct identification of candidate compounds as receptors agonists,  
 CC inverse agonists or partial agonists for use as pharmaceutical agents.  
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in  
 CC the exemplification of the present invention.  
 XX  
 SQ Sequence 500 AA;  
 Query Match 99.7%; Score 2013; DB 21; Length 500;

Best Local Similarity 99.7%; Pred. NO. 4.1e-180;  
Matches 397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESGLLRPAVSEVIVLHNYTKLGRARYQPGAGLRADAVVCLAVCAFIYVLENLAVLLV 60  
DB 103 mesgllrpapvsevivlhyntgklrgayqpgaglradaavclavcafiylenlavllv 162

QY 61 LGRHPRHAPMFLLSGLTSLDLAGAAYAAANLLSGPLTKLSPALWFAREGGVFVALT 120  
DB 163 lgrhprhampmflilsgltslldlagaayaanllsgpltklspalwfareggvfvalt 222

QY 121 ASVLSLLAIALERSLTMAARRGPAPVSSRGRTLAMAAAGVSLLLGLLPALGWNCGLRLD 180  
DB 223 asvlsllaialersltmarrrgpapvssrgtrtlaaaaagvslilglpawncglrld 282

QY 181 ACSTVPLXYAKAVLFCVLAFAVIGLAAICALLYARIYCOVRANARLPARPGTAGTTSTRA 240  
DB 283 acstvpilyakavlfvcvllaafvlglaaicallyariycovranarrrparpgtagttstra 342

QY 241 RRKPRSLALRTLSVLLAFVACWGPFLFLLLLLDVACPARTCPVLLQADPFLGLAMANS 300  
DB 343 rrpksrallrtlsvllafvacwgpflfllllldvacpartcpvllqadpflglamansl 402

QY 301 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPSSQSQSASAAEASGGRLRCLPGLDSEF 360  
DB 403 lnpiiytltnrdlrhallrvccgrhscgrdpssgqgsasaaesggrrlclppldgsf 462

QY 361 SGERSPQRDLTSGTSGSPGAPTAARTLVSEPAAD 398  
DB 463 sgerpspqrldtsgtsgspgaptartlvsepaad 500

RESULT 7  
AAV72561  
ID AAV72561 standard; Protein; 400 AA.  
AC AAV72561;  
XX  
DT 02-MAY-2001 (first entry)  
XX  
DE Rat G protein-coupled receptor (GPCR), B1C3.  
XX  
KW Rat; G protein-coupled receptor; GPCR; B1C3; intracellular signal;  
XX  
OS analgesic; anaesthetic; drug screening.  
XX  
OS Rattus sp.  
XX  
PN WO200105829-A1.  
XX  
PD 25-JAN-2001.  
XX  
PF 20-JUL-2000; 2000WO-SE01505.  
XX  
PR 21-JUL-1999; 99SE-0002763.  
XX  
PA (ASTR ) ASTRAZENECA AB.  
XX  
PI Ahmad S, Hoffert C, O'donnell D, Pelletier M, Walker P;  
XX  
DR WPI; 2001-159515/16.  
DR N-PSDB; AAD02561.  
XX  
PT New G protein-coupled receptor, B1C3, expressed in the central nervous  
PT system of rats, for screening modulators of the receptor and  
PT identifying agents of therapeutic use capable of binding to the  
PT receptor -  
XX  
XX Claim 1; Fig 2; 35pp; English.  
PS  
PS The present sequence is rat G protein-coupled receptor (GPCR), B1C3  
CC which is expressed in the central nervous system of rats.  
CC B1C3 nucleic acids and recombinant proteins are useful in assays to  
CC identify agents capable of binding to the receptor, which modulate

CC intracellular signalling such as adeny cyclase activity or  
CC intracellular calcium concentration. These agents have potential  
CC therapeutic applications as either analgesics or anaesthetics. B1C3  
CC receptors are also useful for screening drug candidates using cell  
CC signalling assays. Antibodies to B1C3 are useful in purification of  
CC intact receptor or fragments of the receptor.  
XX  
SQ Sequence 400 AA;

Query Match 84.9%; Score 1714.5; DB 22; Length 400;  
Best Local Similarity 87.2%; Pred. No. 2.7e-152;  
Matches 346; Conservative 14; Mismatches 34; Indels 3; Gaps 3;

QY 1 MESGLLRPAVSEVIVLHNYTKLGRARYQPGAGLRADAVVCLAVCAFIYVLENLAVLLV 60  
DB 1 mesgllrpapvsevivlhyntgklrgaryqpgaglradaavclavcafiylenlavllv 60

QY 61 LGRHPRHAPMFLLSGLTSLDLAGAAYAAANLLSGPLTKLSPALWFAREGGVFVALT 120  
DB 61 lgrhprhampmflilsgltslldlagaayaanllsgpltklspalwfareggvfvla 120

QY 121 ASVLSLLAIALERSLTMAARRGPAPVSSRGRTLAMAAAGVSLLLGLLPALGWNCGLRLD 180  
DB 121 asvlsllaialersltmarrrgpapvssrgtrtlaaaaagvslilglpawncglrld 180

QY 181 ACSTVPLXYAKAVLFCVLAFAVIGLAAICALLYARIYCOVRANARLPARPGTAGTTSTR 239  
DB 181 acstvpilyakavlfvcvllaafvlglaaicallyariycovranarrrlraggsrratssr 240

QY 240 AFRKPRSLALRTLSVLLAFVACWGPFLFLLLLLDVACPARTCPVLLQADPFLGLAMANS 299  
DB 241 srhtrpsrallrtlsvllafvacwgpflfllllldvacpartcpvllqadpflglamans 300

QY 300 LNPIIYTLNRDLRHALLRVCCGRHSCGRDPSSG-QQSASAAEAS-GGLRRLCLPPGLD 357  
DB 301 lnpiiytltnrdlrhallrvccgrhscgrdpssnslqrspsavgpsgglrrclppltd 360

QY 358 SFSGSSERSRQRDLTSGTSGSPGAPTAARTLVSE 394  
DB 361 rsspsehsqprdgmdtscstgspgaatanrtlvpd 397

RESULT 8  
AAV05489  
ID AAV05489 standard; Protein; 382 AA.  
XX  
AC AAV05489;  
XX  
DT 07-JUL-1999 (first entry)  
XX  
DE Human EDG-2 protein sequence.  
XX  
KW EDG-1; EDG-2; EDG-3; EDG-4; EDG-5; PSP-24; human; detection; therapy;  
KW inverse agonist; allosteric modulator; lysophosphatidic acid receptor;  
KW LPA signalling mediated disease; cellular apoptosis.  
XX  
OS Homo sapiens.  
XX  
PN WO9919513-A2.  
XX  
PD 22-APR-1999.  
XX  
PF 09-OCT-1998; 98WO-US21315.  
XX  
PR 10-OCT-1997; 97US-0061572.  
XX  
PA (LXRB-) LXR BIOTECHNOLOGY INC.  
XX  
XX Erikson J, Goddard JC, Kiefer M;  
XX WPI; 1999-277658/23.  
DR N-PSDB; AAX36566.  
DR

XX	Identification of (antagonists of LPA receptor EDG-2 for, e.g.
PT	treating LPA signalling mediated diseases such as cellular apoptosis
XX	
XX	Claim 1; Page 51-52; 63pp; English.
PS	
XX	
CC	This sequence represents human EDG-2.
CC	The invention relates to methods of detecting (antagonist, inverse
CC	agonist or allosteric modulators of the lysophosphatidic acid receptors
CC	EDG-1, EDG-2, EDG-3, EDG-4, EDG-5, and PSP-24. The methods are used to
CC	identify (antagonists and allosteric modulator of the lysophosphatidic
CC	acid (LPA) EDG2 receptor, e.g. to treat LPA signalling mediated disease
CC	such cellular apoptosis.
XX	
XX	Sequence 382 AA; SQ

```

Query Match      42.9%; Score 865.5; DB 20; Length 382;
Best Local Similarity 48.5%; Pred. No. 7.6e-73;
Matches 181; Conservative 66; Mismatches 105; Indels 21; Gaps 6;

QY 13 EVIVLHYNTKLR-GARYOPGAGLRADAVVCLVACAFIVLENLAVLLVGLGRHPFAPHM 71
Db 23 diivrhnytklnlga--ekdhgikltsvfllccfillenfvlltwtkxfhrpm 80

QY 72 FLLGLSLTSDLLAGAAAYANILLSPGLTKLSPALWFAREGGVFALTASVLLSLALAI 131
Db 81 yyfignalsdlagvaytanillsgattkllpaqwfiregsmfvalsasvllalal 140

QY 132 ERSLLPMWRGAPAPVSSRGRLTMAAAANGVSLGLLPLALGNCLGRIDACSTVPLVAK 191
Db 141 eryllmkmklhngsnrsrflisacwslilgplmgwncisslscstvplyhk 200

QY 192 AYVLECVLAFVGIILAAICAIYARIYCOVRANARLRPARPCTAGTSTARRKPRSLALLR 251
Db 201 hylifcttvfllllstailcryslvrtrsrllfrk----nlskgrsseklsallk 256

QY 252 TUSVLLAFVACWGPFLFLLLLVLOVADPARTCPVLLQADPFLGLAMANSLLNPYYLTNR 311
Db 257 twilvslvfiacwaplfilllldvgckactcdillykaeyflvlnsgtnpilytlcnk 316

QY 312 DURHALLRLVCGRRGCRDPSSGQSQAASAEAGGLRRCLEPGLDGSGFSGSRSS-PQR 370
Db 317 emtrafrilv-----sckcpng-----dsagkfkrplipgmefrsksdnshpqk 363

QY 371 DGLDTSGTGSPPG 383
Db 364 dgdgdnpetimssg 376

```

RESULT 9  
AAU00302  
ID AAU00302 standard; Protein; 382 AA.

18-AUG-1993; 99US-0376399.  
(ATAI-) ATAIRGIN TECHNOLOGIES INC.  
Erickson J, Goddard JG, Kiefer M, Picker D;  
WPI; 2001-226550/23.  
DR DR N-PSDB; RA500257.  
Modulating activity of a lysophosphatidic acid or its receptor for  
treating tumours, viral diseases, involves introducing to LPA or its  
receptor a composition containing diol, epoxide or phosphate compound  
Disclosure: Page 76-78; 89pp; English.

The sequence represents the LPA receptor-related amino acid sequence #1 used during analysis of lysophosphatidic acid receptor (LPA), EDG-2. The sequence is given in the specification but no further information is given. LPA is a phospholipid found in a variety of plant and animal products. EDG-2 is involved in cell signalling through activation of a MAP kinase cascade-dependent reporter. Modulating the activity of a lysophosphatidic acid (LPA) or LPA receptor (EDG-2) involves introducing a composition comprising LPA modulators to the LPA or receptor. The method is useful for treating diseases characterised by slowed growth or repair of neuronal cells, neurodegenerative diseases, such as Alzheimer's disease, Parkinson's disease, and acute neuron damage, for modulating apoptotic pathways and treating ischaemic heart disease, tumours, viral diseases bowel disease, and rejection of organ transplants.

Sequence 382 AA:

```

Query Match          42.9%; Score 865.5; DB 22; Length 382;
Best Local Similarity 48.5%; Pred. No. 7.6e-73;
Matches 181; Conservative 66; Mismatches 105; Indels 21; Gaps 6;

13 EVIVLHYNTKLR-GARYOPGAGLRADAVVCLAVCAFTVLENLAVLLVLRHPRFHAPM 71
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
23 diivrhnytdklniga--ekdhgiktsvfillccfillenifvllciwtkkfrpm 80
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
72 FLLGSLTLDLLAGAAYANILSGPLTLKSPALWFAREGGVFVALTASLLSLAL 131
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
81 yyifgnlaalsdlagvaytanlllegattyklcpaqwfiregsmfvalsasvfallai 140
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
132 ERSITWARRGAPVSSRGRTLMAAAANGVSLLLGLPALGHNCILGRLDACSIVLYPAK 191
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
141 erylunlkmklnhgnsrsrflfiascwlslilgplbmgnwncisslscstvlplyhk 200
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
192 AYVLFVCLAFYGIILAAICALYARIYCOVRANARRLPARPCTAGTSTARRRRPSLALLR 251
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
201 hylftttvftlllsiallycrylsivtrsrlyfrk-----niskgrasekalalk 256
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
252 TLSVVFLAFVACWGPFLFLLLDLVACPARTCPVLLQADPFLGLMANSLNPIITVLTNR 311
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
257 tvliivslviacwaplfillllldvgckactcdilykaeyflwlvinsgtnpilytlcnk 316
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
312 DLRIALLRVCCRGSRDPSPGSQOSASAAEASGLRCLPPGLDGGFGSGSERSS--PQR 370
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
317 emrrafrirv-----sckcpng-----dsagkfrplipgmefrsksdnshpqk 363
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
371 DGLDTSGSTGSPG 383
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
364 dgdndpnetimssg 376
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 10  
 AAB28521  
 ID AAB28521 standard; Protein; 382 AA.  
 XX  
 XX AAB28521;  
 XX AC



06-FEB-2001 (first entry)	XX	Mouse EDG1 polypeptide.	XX	252	TLISVLLAFVACMGPFLLLLLDVACPAPARTCPVLLQADPFLGAMNSLLNPIIYTLNR	311
Mouse EDG1 polypeptide.	XX		XX	257	TVIIIVSVFIACWAPFIILLIDVGCKAKTCDILYKAEYFLVAVINSQNPILYTLTK	316
Mouse; EDG1; antimicrobial; cytostatic; analgesics; antidiabetic; antiparkinsonian; cardiant; osteopathic; anorectic; antiasthmatic; antiemetic; tranquilizer; antiallergic; neuroleptic; antiulcer; antidepressant; infection; cancer; diabetes; eating disorder; asthma; Parkinson's disease; heart failure; hypotension; hypertension; osteoporosis; angina pectoris; stroke; ulcer; allergy; benign prostatic hypertrophy; migraine; vomiting; neurological disorder.	XX		XX	312	DLRHALLRLVCCGHRHSCGRDPSCSQSASNAEASGGLRCLPPGLDGSFGSERS-POR	370
benign prostatic hypertrophy; migraine; vomiting; neurological disorder.	XX		XX	317	EMIRAFIRIV-----SCCKCPNG-----DSAGKFKRPIIPGMFESRSKDSNASHPQK	363
Mus musculus.	OS		XX	371	DGLDTSGTSGSPG	383
WO200059529-A1.	PN		XX	364	dGdnpetlmssg	376
12-OCT-2000.	PD		XX	AAW01664	standard; Protein; 383 AA.	
05-APR-2000; 2000WO-US09075.	PF		XX	AAW01664;		
05-APR-1999; 99US-0127696.	PR		XX	01-APR-1997 (first entry)		
04-APR-2000; 2000US-0127696.	PR		XX	p(rat-edg), G-protein coupled receptor.		
(SMIK ) SMITHKLINE BEECHAM CORP.	FA		XX	p(H218); G-protein coupled receptor; cell differentiation; proliferation; proline directed kinase; cell division; growth factor response; rat-edg; therapy; diagnosis.		
Tau1 P;	PI		XX	Rattus rattus.		
WPI; 2000-664960/64.	DR		XX	US5585476-A.		
N-PSDB; AAC63385.	DR		XX	17-DEC-1996.		
Murine EDG1 polypeptide useful for treating infections and diseases such as cancer, diabetes, Parkinson's disease, allergies, and neurological disorders such as anxiety schizophrenia, depression	PT		XX	15-FEB-1994; 94US-0196989.		
Claim 1; Page 31-32; 34pp; English.	PS		XX	15-FEB-1994; 94US-0196989.		
The present sequence is mouse EDG1. Agonists and antagonists of mouse EDG1 are useful for treating patients with enhanced or inhibited activity or expression of EDG1. The EDG1 polypeptide is useful for treating abnormal conditions such as infections caused by bacteria, fungi etc., pain, cancer, diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, myocardial infarction, stroke, ulcers, allergies, benign prostatic hypertrophy, migraine and vomiting. It is useful for treating psychotic and neurological disorders including anxiety, schizophrenia, manic depression, depression, delirium, dementia and severe mental retardation and dyskinesias such as Huntington's disease or Gilles de la Tourette's syndrome.	CC		XX	DNA encoding rat protein p(H218) - associated with cell proliferation and/or differentiation		
Sequence 382 AA;	XX		XX	Example 9; Column 17-20; 33pp; English.		
Query Match 42.8%; Score 863.5; DB 21; Length 382;	XX		XX	This sequence is a novel rat protein p(rat-edg), a member of the G-protein coupled receptor (GPR) superfamily of proteins. The amino acid similarity between p(H218) (AAW01663) and p(rat-edg) suggests that they may be activated by the same endogenous ligand(s). The expression pattern of mRNA transcripts of both genes in cell lines, various rat tissues and developing rat brain suggests that they both play a role in cell proliferation and/or differentiation.		
Best Local Similarity 48.5%; Pred. No. 1.2e-72;	XX		XX	Sequence 383 AA;		
Matches 181; Conservative 65; Mismatches 106; Indels 21; Gaps 6;	XX		XX	Query Match 42.3%; Score 853.5; DB 18; Length 383;		
	XX		XX	Best Local Similarity 47.6%; Pred. No. 1e-71;		
	XX		XX	Matches 177; Conservative 66; Mismatches 110; Indels 19; Gaps 5;		
	XX		XX	13 EVTLVHNYTGKLR-GARYQPGAGLRADAVYCLAVCAFIIVLENLVLGRHPRHPM 71		
	XX		XX	23 diivrhnytgklniga--ekdhgikitsvfvilcofillenifvltiwtkkfhrpm 80		
	XX		XX	72 FLLGSLTLDLAGAAYANILLSPGLTKLSPALWFAREGVFVALTSVLSLALAI 131		
	XX		XX	81 yfignalsldilagvaytanllsgattkltpaqvfregmfvalsasvfallalal 140		
	XX		XX	132 ERSLLWARGPAPVSSRGRTLAMAAAWGVSLLGLLPALGWNCGLGRDACSTVPLPYAK 191		
	XX		XX	141 erytalmkmlhngnsersflliscacwvslilgipimgncislsascstviplyhk 200		
	XX		XX	192 AYVLCVLAFAVGILAAICALYARIYCOVRANRRLPARPCTAGTTSTRARRPRSLALLR 251		
	XX		XX	201 hyliftctvtllslslvlycrlsylvtrsrlltrfk-----nisksrsekslallk 256		

```

Db 143 ryitlmkmlhngnsr:ffllisacwvsiilggplmgwncisslscstviplyhkh 202
QY 193 YVLCVLAFAVGILAAICALYARIYCOVRANARRLPARGTAGTSTRARRKPRSLALLRT 252
Db 203 yilfctvtllllsivlycrylsrvtrrrlfrk-----niskasrseksalllkt 258
QY 253 LSVLLAFVACWGLFLLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRD 312
Db 259 viilsvfiacwapiifillldvgckaktcdilykaeyflvlnsgtnpilytltnke 318
QY 313 LRHALLRLVCCRHSCGRDPGSGSQSASAAEASGLRLRCLPPGLDGSFSGSERS-PORD 371
Db 319 mrrafirii-----sckcpng-----dsagkfrpiipgmefsrsksdnshpqkd 365
QY 372 GLDTSSTGSPG 383
Db 366 dgdnpetlmssg 377

RESULT 12
AAW87791
ID AAW87791 standard; Protein; 383 AA.
XX
AC AAW87791;
XX
DT 12-MAR-1999 (first entry)
XX
DE Rat-edg, G-protein coupled receptor superfamily member.
XX
KW H218; G-protein coupled receptor superfamily; cell proliferation;
KW cell differentiation; cancer; rat-edg; diagnosis; therapy.
XX
OS Rattus sp.
XX
PN US5856443-A.
XX
PD 05-JAN-1999.
XX
PF 06-DEC-1996; 96US-0760936.
XX
PR 15-FEB-1994; 94US-0196989.
XX
PR 06-DEC-1996; 96US-0760936.
XX
PA (MACL/) MACLENNAN A J.
XX
PI MacLennan AJ;
XX
DR WPI; 1999-105192/09.
XX
DR N-PSDB; AAV84036.
XX
PT New isolated polynucleotide encoding a G-protein coupled receptor -
PT used to develop products for use in diagnosis and therapy of
PT conditions involving cell proliferation and differentiation
XX
PS Disclosure; Columns 23-26; 33pp; English.
XX
CC The present sequence represents a protein designated rat-edg. The rat-edg
CC polypeptide is a member of the G-protein coupled receptor superfamily.
CC The specification also describes a protein designated H218, which
CC is also a member of the G-protein coupled receptor superfamily. H218
CC is involved in cell proliferation and differentiation, and in disease
CC states such as cancer. The polynucleotides and polypeptides can be
CC used to develop products for diagnosis and therapy.
XX
SQ Sequence 383 AA;

Query Match 42.3%; Score 853.5; DB 20; Length 383;
Best Local Similarity 47.8%; Pred. No. 1e-71;
Matches 177; Conservative 66; Mismatches 110; Indels 19; Gaps 5;

QY 13 EVIVLHNYTGKLRGARYQPGAGLRADAVVCLAVCAFIIVLENLVLVLRHPRHAPMF 72

```

```

Db 24 diivhnytytkl-nlgvekhgikltsvffilicollifenfvlltwktkkfrmpy 82
QY 73 LLGSLTSLDLLAGAAYAAANTLLSGPLTKLSPALWFAREGGVFVALTSVLSLLAIALE 132
Db 83 yfignalaedilagvaytanillsgattykltpaqwflregsmfvalsaavfallaiaie 142
QY 133 RSLTWARGPAPVSSRGFTLMAAANGVSLLLGLLPALGNCLGRDACSSTVPLIYAKA 192
Db 143 ryitlmkmlhngnsr:ffllisacwvsiilggplmgwncisslscstviplyhkh 202
QY 193 YVLCVLAFAVGILAAICALYARIYCOVRANARRLPARGTAGTSTRARRKPRSLALLRT 252
Db 203 yilfctvtllllsivlycrylsrvtrrrlfrk-----niskasrseksalllkt 258
QY 253 LSVLLAFVACWGLFLLLLLDVACPARTCPVLLQADPFLGLAMANSLLNPIIYTLNRD 312
Db 259 viilsvfiacwapiifillldvgckaktcdilykaeyflvlnsgtnpilytltnke 318
QY 313 LRHALLRLVCCRHSCGRDPGSGSQSASAAEASGLRLRCLPPGLDGSFSGSERS-PORD 371
Db 319 mrrafirii-----sckcpng-----dsagkfrpiipgmefsrsksdnshpqkd 365
QY 372 GLDTSSTGSPG 383
Db 366 dgdnpetlmssg 377

RESULT 13
AAW33205
ID AAW33205 standard; Protein; 381 AA.
XX
AC AAW33205;
XX
DT 18-NOV-1999 (first entry)
XX
DE Human EDG-1c receptor protein.
XX
KW EGD-1c receptor; sphingosine-1-phosphate; S-1-P; human; anti-arrhythmic;
KW dihydro sphingosine-1-phosphate; anti-HIV; anti-asthmatic; cytostatic;
KW anti-anginal; analgesic; antimicrobial; neuroprotective; nootropic; pain;
KW treatment; infection; bacterial; fungal; protozoan; viral; HIV-1; HIV-2;
KW cancer; diabetes; obesity; anorexia; bulimia; Parkinson's disease;
KW asthma; acute heart failure; hypertension; hypotension; osteoporosis;
KW urinary retention; angina pectoris; myocardial infarction; stroke;
KW congestive heart failure; left ventricular hypertrophy; arrhythmia;
KW restenosis; coronary artery angioplasty; vascular sclerosis; allergy;
KW inflammation; deleterious fibrosis; atherosclerosis; angiogenesis;
KW wound healing; ulcer; benign prostatic hypertrophy; migraine; vomiting;
KW psychotic disorder; neurological disorder; anxiety; schizophrenia;
KW manic depression; depression; delirium; dementia; degenerative disease;
KW severe mental retardation; neurodegenerative disease; ischemic stroke;
KW dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome.
XX
OS Homo sapiens.
XX
PN WO9946277-A1.
XX
PD 16-SEP-1999.
XX
PF 04-MAR-1999; 99WO-US04675.
XX
PR 09-MAR-1998; 98US-0077369.
XX
PR 28-MAY-1998; 98US-0087102.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PA (SMIK ) SB LAB PHARM.
PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX (STAD/) STADEL P.
PI Bergsma DJ, Chambers JK, Chan W, Johnson RK, Khandoudi N, Livi GP;
PI Robert P, Stadel JM, Wilson S;
XX

```





